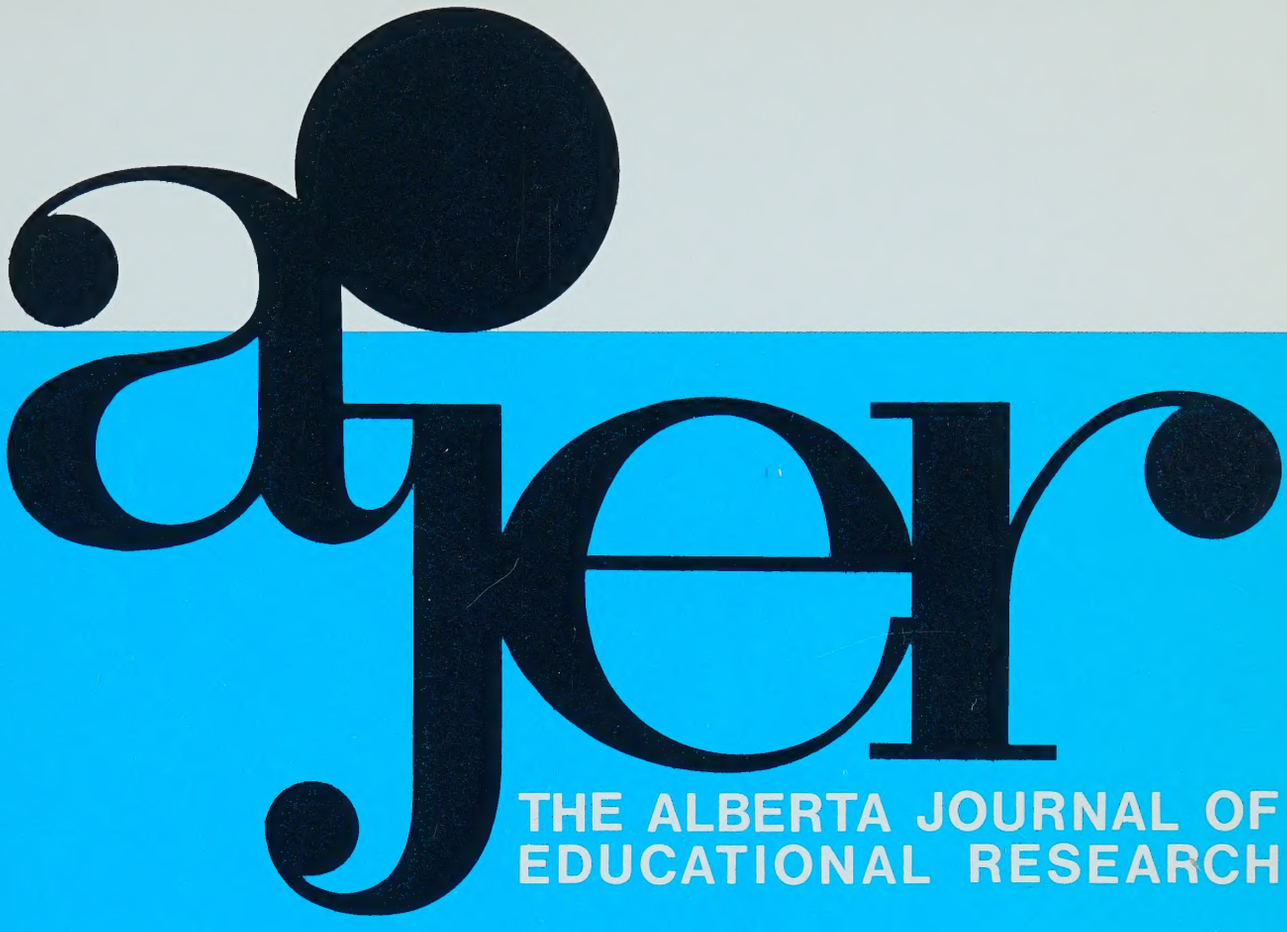

THE ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

VOLUME XXV NUMBER 4 DECEMBER 1979

PUBLISHED BY
THE UNIVERSITY OF ALBERTA • EDMONTON

THE ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

A quarterly journal devoted to the dissemination, criticism, interpretation and encouragement of all forms of systematic enquiry into education and fields related to or associated with education.

Published quarterly in March, June, September, December by the
Faculty of Education, The University of Alberta

CONSULTING EDITORS

J. Britton
University of London

J. Calam
The University of British Columbia

M. Connelly
The Ontario Institute for Studies in Education

K. De Clerck
State University of Ghent

R. N. Evans
University of Illinois at Urbana-Champaign

R. H. Farquhar
University of Saskatchewan

E. Gagné
University of Ottawa

G. Harman
University of Melbourne, Australia

S. Hunka
The University of Alberta

J. W. G. Ivany
Simon Fraser University

D. A. MacIver
University of New Brunswick

L. D. Nelson
The University of Alberta

W. C. Nesbit
Memorial University of Newfoundland

E. Pedersen
McGill University

EDITOR: H. W. Hodysh

SECRETARY: A. Onishenko

FACULTY PUBLICATIONS COMMITTEE

M. A. Assheton-Smith
T. P. Atkinson
N. C. Bhattacharya
R. C. Bryce
W. T. Fagan
H. W. Hodysh
R. G. Martin (Chairman)
E. Miklos
V. R. Nyberg
C. H. Preitz

Editorial policy and the discussion and disposition of manuscripts are the joint responsibility of the Publications Committee. The views expressed and the accuracy of the statements made are the responsibility of the individual authors. The editor is solely responsible for the editorial comments.

AJER gratefully acknowledges support from the Social Sciences and Humanities Research Council of Canada and from the Alberta Advisory Committee for Educational Studies.

AJER is indexed in the *Canadian Education Index, Current Contents/Social and Behavioral Sciences,* and *Social Science Citation Index;* appropriate articles are abstracted in *Educational Administration Abstracts, Psychological Abstracts, Sociology of Education Abstracts,* and *Language Behavior Abstracts.*

The subscription rate is $8.00 per year; single copies are $2.50 each. Please make cheques payable to *The Alberta Journal of Educational Research.* All back issues are available; rates supplied on request. Claims for undelivered copies must be received within three months of the month of publication.

Address all communications and manuscript submissions to the Editor, *The Alberta Journal of Educational Research,* Faculty of Education, 732 Education South, The University of Alberta, Edmonton, Canada, T6G 2G5.

SECOND CLASS MAIL REGISTRATION NUMBER 1436

ISSN 0002-4805

The Alberta Journal of Educational Research

Volume XXV, Number 4 *December, 1979*

CONTENTS

Heuristic Strategies Utilized by High School Students 213
L. Pereira-Mendoza

Newfoundland Dialect Interference in Fourth Grade Spelling 221
L. Walker

The Development in Children and Adolescents of the Construct of Rational Numbers as Operators 234
T. E. Kieren and B. Southwell

The Relationships of Language Concepts, Auditory Comprehension, Visual Perception, and Spatial Relations as Predictors to Reading Achievement in First Grade 248
B. B. Adams and L. O. Ollila

Comparative Difficulty of Beginning Reading Vocabulary: Set II 259
B. Horodezky

Teacher Socialization: The First Five Years 264
J. McArthur

Student Teacher Telephone Conferencing with Satellite Maps as a Monitoring Device 275
J. M. Kirman and J. Goldberg

PERSPECTIVES

Some Observations Regarding Value Added by Education 284
R. N. Evans

ESSAY REVIEWS

Education in Atlantic Canada 288
D. A. MacIver

Moral Education: The Risk of Oversimplification 294
I. DeFaveri

(over)

BOOK REVIEWS
Classroom Language: What Sort?, J. Richards 307
Reviewed by Ellen M. Regan
An Experiment in Teacher Education,
A. Ross, D. McNamara and J. Whittaker 309
Reviewed by Lorne D. Stewart

TABLE OF CONTENTS, VOLUME XXV, 1979 310

FACULTY OF EDUCATION
The University of Alberta

The Alberta Journal of Educational Research *Vol. XXV, No. 4, December, 1979*

LIONEL PEREIRA-MENDOZA
Memorial University of Newfoundland

Heuristic Strategies Utilized by High School Students

This study investigated the heuristic strategies utilized by high school students (N = 20) to solve mathematical problems. The experimental session lasted up to one hour per subject, during which time the student attempted four problems, a maximum of fifteen minutes being allowed per problem. The subjects verbalized their ideas while solving the problems and the sessions were videotaped. Analysis of the protocols employed by the subjects indicated that examination of cases (including systematic cases and looking for a pattern in the results), symmetry, and analysis were the most common heuristics. Subjects tended to examine cases in many problems while employing the other heuristics in only one problem, indicating that generally the utilization of heuristics tended to be problem-dependent.

In a world in which scientific and technological knowledge is growing exponentially, society is presented with the challenge of educating individuals who will be able to function in a continually changing environment. Thus, educators and psychologists are faced with the problem of identifying and teaching the competencies needed to deal with this changing world. At the center of these competencies are "strategies of inquiry." Heathers (1965) indicated that "problem-solving thinking or inquiry is generally considered to be the core of the educational process, and the chief mark of the educated person" (p. 2).

One approach taken by scholars has been to investigate the nature of the problem-solving process. Their efforts have resulted in analysis of various stages involved in problem-solving. For example, Bloom and Broder (1950) suggested the following four components: understanding the nature of the problem, understanding the ideas contained in the problem, general approach to the solution of problems, and attitude towards the solution of

---

Dr. Pereira-Mendoza is an associate professor of mathematics education at Memorial University, St. John's, Newfoundland, Canada. He received his undergraduate education in England and his doctoral work was undertaken at the University of British Columbia. His primary research interest is in the area of mathematical problem-solving, with particular reference to the heuristics utilized in solving novel problems.

problems. Garry and Kingsley (1970, p. 464) suggested three steps in problem-solving: the search phase, the functional solution phase, and verification of the final solution. Gagné (1966, p. 138) included the stages of defining the problem by distinguishing essential features, searching for and formulating hypotheses, and verifying the solution in his analysis of the problem-solving process. While such analyses provide an overall framework for examining the nature of the problem-solving process, they do not include specific techniques that can be applied by the learner to make discoveries about novel mathematical problems. These specific techniques can be defined as heuristics.

Mathematical heuristics are procedures which are used for the purpose of discovering mathematical relationships in a novel problem, but whose application does not guarantee success. For the mathematics educator, the work of Polya (1954a, 1954b, 1957, 1962, 1965) has provided the major source of heuristic strategies. Polya's heuristic framework is stated in the form of specific questions that can be posed in order to try to solve problems. Some of these questions are: "Do you know a related problem? . . . Could you imagine a more accessible related problem? A more general problem? A more special problem? An analogous problem? Could you solve part of the problem? Draw a figure . . ." (Polya, 1957, pp. xvi-xvii). MacPherson (see Blake, 1976) refined the Polya framework to produce a set of heuristics, and it is this set that formed the framework for this research. Some of the heuristics in this set are presented under a subsequent heading, *Coding Form.*

Research and Purpose

Much of the current interest in clinical research on heuristic strategies in the area of mathematics education was initially generated by Kilpatrick's study (1967) which utilized a coding scheme based on Polya's heuristic framework. While other studies (e.g., Lucas, 1972; Webb, 1975; Kantowski, 1977) tended to employ distinct but not disjoint sets of heuristic strategies and coding schemes, it seems reasonable to conclude that most researchers utilized the Polya framework and Kilpatrick's coding scheme as the initial basis for their research, each researcher modifying them in view of additional research findings. The overall implications of such research seem to be clear, namely that students do utilize and can be taught to utilize various heuristic strategies. For example, in many of these studies, it was found that students drew figures, made tables, looked for patterns, and tried trial and error (random cases) when solving problems. In these studies, algebraic or word problems were often employed to generate heuristic strategies.

The purpose of this study is to utilize the heuristics outlined by MacPherson as a basis for investigating the strategies employed by students in solving geometric problems. Furthermore, this study allows the students to utilize physical materials when solving problems, a departure from the other studies. The specific questions investigated were:

1. Without specific training in heuristics, what heuristic strategies are employed by high school students when solving novel mathematical problems?
2. What are the most commonly utilized heuristics?

Procedure

A sample consisting of 20 high school students (ten in grade 9 and ten in grade 11) was selected from the local schools. The subjects were all in the honours mathematics program and had been identified by their teachers as the better problem solvers.

Each subject came to the university at a different time and was taken to a room where the videotaping equipment had been set up. Subjects were given a general orientation towards the equipment and an opportunity to ask any general questions regarding the taping. It was explained that they were expected to think aloud while solving the problems. Finally, the subjects were informed of the maximum of 15 minutes per problem, and that at the end of this period or on completion of the problem (whichever came first), they would continue on to the next problem. It should be noted that a problem was considered complete when the subject obtained a solution, whether or not that solution was correct. No comments or advice were given during the videotaping, except for the occasional reminder to think aloud.

Problems

The four problems are listed below. In addition to the materials appropriate to each problem, subjects were provided with writing paper, pens, and a ruler.

The Chinese Tangram Puzzle. The seven pieces were placed in an envelope in front of the subject, together with an index card on which was written:

Arrange all the pieces in the envelope to form a square.

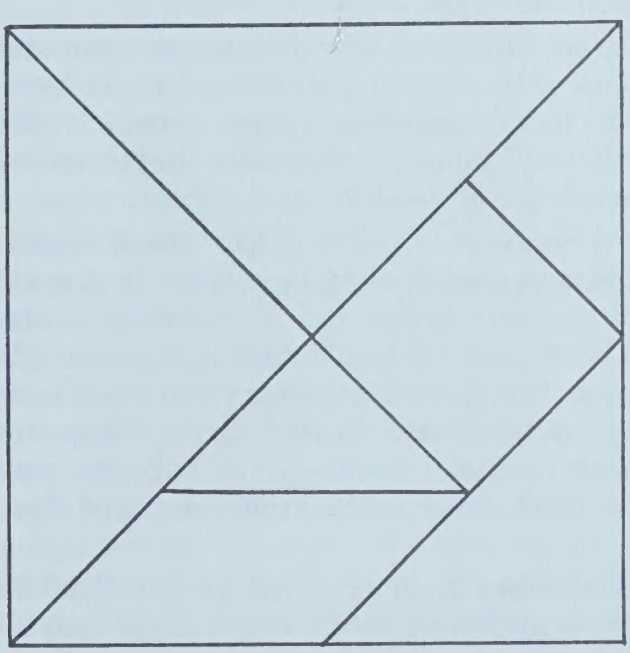

Figure 1. Tangram solution.

1,000 Squares Problem. The subject was presented with a plastic bucket containing a large number of congruent squares, together with an index card on which was written:

Given 1,000 congruent squares how can they be arranged in a plane so that the total perimeter of the figure formed is a minimum? Find the perimeter.

Ten Lines Problem. The subject was presented with a geoboard, elastic bands, and an index card on which was written:

Find the maximum number of points of intersection of ten lines in a plane.

5 x 5 Rectangle Problem. This problem was an adaptation based on the classic problem: How many squares on a chessboard? The subject was presented with a ruled piece of card (approximately 25 cm by 25 cm) divided into 25 congruent squares, together with an index card on which was written:

Find the total number of rectangles on this board.

Coding Form

The coding form employed was basically that developed by Blake (1976); a detailed discussion can be found in his dissertation. Since this study was concerned with the heuristics utilized by students, only that aspect of the form is discussed. The coding form enabled the coder to enter both heuristics employed and the order in which they were applied. The form contained twelve heuristics including: *smoothing*—when the problem is altered in order to obtain some isomorphism between the problem and a mathematical system; *analysis*—breaking the problem into subproblems; *cases*—renaming the variables in a problem as constants (includes all cases, critical and systematic cases); *exclusion*—a variable is excluded from the problem; and *symmetry*—making use of intrinsic or constructed symmetry in a problem. A detailed account of the twelve heuristics, together with examples, can be found in Blake (1976, pp. 7-18).

Analysis of the Data

Using videotapes retained from the pilot study, the judges were trained in the coding procedure. At the end of the training sessions, the analysis of the data tapes was undertaken. The judges coded each tape separately, then met to discuss the results. Only when all judges agreed that there was evidence of a given strategy was that evidence utilized in the analysis.

The most commonly utilized heuristic was one of the "cases family." It is important to note that while some subjects examined cases in a systematic manner, not all of them used the data to try and formulate a pattern. For example, in the Ten Lines problem, some subjects tried 2, 3, 4, and 5 lines in a systematic manner (obtaining the correct answers), but then concluded that the problem was getting complex and started again. They made no attempt to search for a pattern in their results. Some subjects started to search for a pattern after only three cases while others required five or six cases.

The two examples of critical cases both occurred in the 1,000 Squares problem. In each case, the subject made comments to the effect that if all the squares were separated, the perimeter would be 4,000 and hence the answer must be less than that. Thus, the subjects utilized this as a basis for establishing a bound for the answer.

Attempts to solve the problems employing symmetry occurred most

TABLE 1
USE OF HEURISTICS[a]

Heuristic	No. of subjects who utilized the heuristic at least once when solving the four problems (Maximum 20)
Cases	14 (utilized at least one of the various facets of this heuristic--see below)
(a) All	4
(b) Critical	2
(c) Systematic	10
(d) Systematic Plus Pattern	6
Analysis	6
Symmetry	8
Analogy	1
Inverse Deduction	1

[a] Note that the use of random cases is not included in this list. The nature of the problems, such as the tangram problem, meant that the subject either had to do nothing or manipulate the equipment in some manner. As such, it was felt that random manipulation could not be considered an heuristic, but rather a function of these specific problems. In other problem situations it could be classified as an heuristic.

frequently in the Tangram and Ten Lines problem (see Figure 2). For example, in the Ten Lines problem, the subjects would *continually* try to form the square by organizing the pieces in a symmetrical manner. It is significant to note that most of the subjects who utilized symmetry continued with a fixation effect on this heuristic. For example, once having decided that the solution to the Ten Lines problem was symmetrical, they would continue with a symmetrical approach, usually resulting in the conclusion that the maximum number of points of intersection was 25. Symmetry was also noted in the 5 x 5 Rectangle problem where a subject would count the 2 x 3 rectangles, say, horizontally and conclude that there were the same number vertically.

The only other heuristic that occurred with any regularity was that of analysis. This heuristic was employed most frequently in the Tangram and 5 x 5 Rectangle problems. In the Tangram problem, the subjects said that they could solve the problem by making two rectangles, or two congruent right triangles, or in one case, four congruent right triangles. They then proceeded to try to manufacture the appropriate figures. In the 5 x 5 Rectangle problem, the subjects tried to solve the problem by breaking it into two subproblems, namely the number of squares and the number of rectangles that were not square. In order for analysis to be coded it was necessary that a subject had *both* made a comment indicating a breaking of the problems into subproblems and followed through by applying the approach suggested by the analysis to try to solve the problem.

It is important to note that a subject who applied an heuristic to one particular problem did not always utilize that heuristic in another problem.

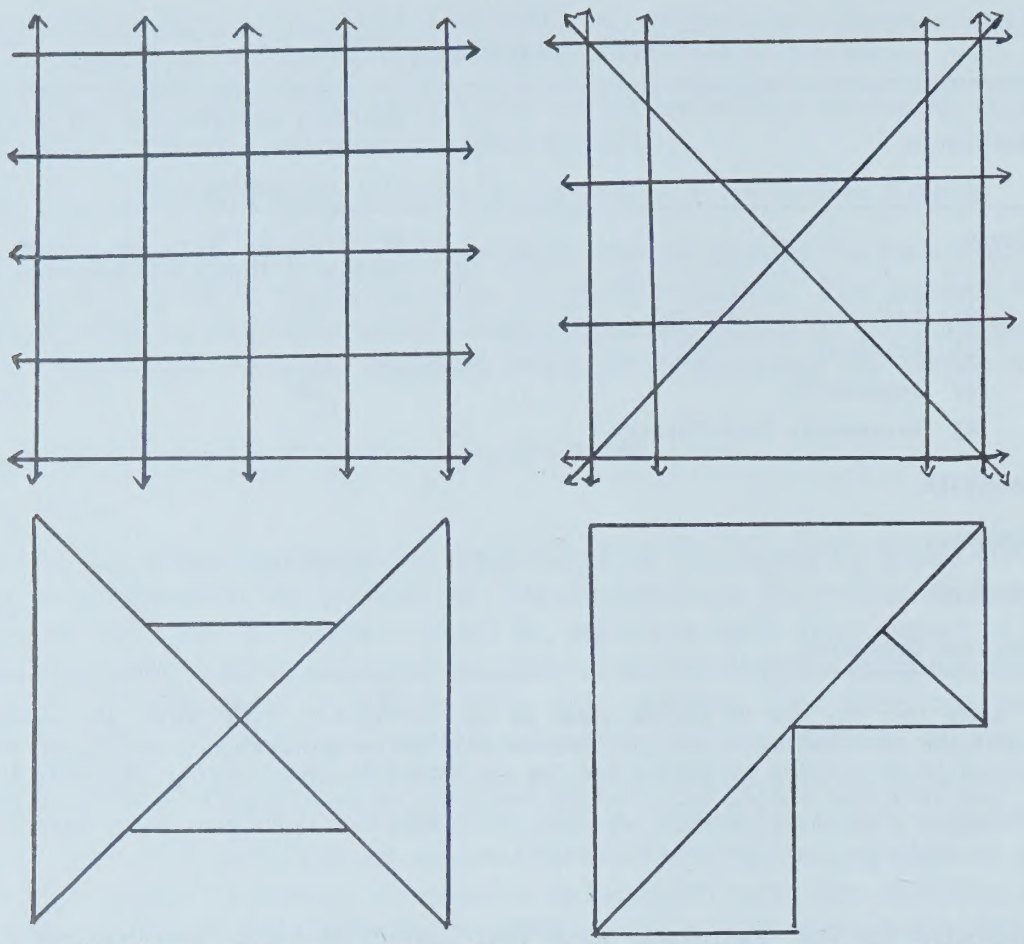

Figure 2. Symmetric arrangement for the Tangram and Ten Lines Problem.

For example, a subject who tried symmetry in the Ten Lines problem did not necessarily try the heuristic when solving the Tangram problem. It appeared that the utilization of an heuristic was *very* dependent on a subject's perception of the problem. This result is consistent with Webb's (1975) conclusion that heuristics were problem-dependent.

Of the 20 subjects involved in the study, 16 applied at least one heuristic during the experimental session. The problem-solving approach utilized by the other four can best be described as totally random "trial and error." Most subjects who applied systematic cases (either with or without searching for a pattern) did so in all the problems, except for the Tangram problem. It appeared that the subjects were willing to try this heuristic in a variety of problem settings. The same conclusion cannot be drawn regarding analysis and symmetry. For example, although symmetry was utilized by a total of 8 subjects, only 2 of them applied the heuristic to two problems.

Table 2 indicates that only three subjects applied three heuristics during their experimental sessions; ten applied two heuristics, usually one of the cases family and one other heuristic. Of the three subjects who utilized only one heuristic, two applied cases and the other applied symmetry,

Finally, one particular aspect of the problem-solving process not discussed so far is the role of physical materials. Analysis of the videotapes indicated that the grade 9 subjects utilized the physical materials far more than the grade 11 subjects. For example, in the Ten Lines problem, *all* the

TABLE 2
NUMBER OF DIFFERENT HEURISTICS UTILIZED BY STUDENTS[a]

	Number of heuristics			
	3	2	1	0
Grade 9	1	4	2	3
Grade 11	2	6	1	1

[a] Subjects who utilized one or more of the components of the cases family have been given credit for one heuristic.

grade 9 subjects used the geoboard, while five of the grade 11 subjects immediately pushed the board aside and started drawing figures (two of the five who pushed the board aside decided to use it later in the problem). Only the Tangram problem produced no differences, with all subjects manipulating the pieces.

Summary of Results

Evidence of five heuristics was found, namely cases (all, critical, and systematic—both with and without searching for a pattern), analysis, analogy, symmetry, and inverse deduction. However, only one example of each of analogy and inverse deduction was found. The most commonly utilized heuristic was cases, particularly systematic cases, followed by symmetry and analysis. No subjects utilized more than three heuristics during the experimental session, with most subjects (10) employing two heuristics. Finally, the grade 9 subjects relied more heavily on the physical materials than the grade 11 subjects, although both groups utilized the materials.

There are many possible directions for further research resulting from this study. First, as was indicated earlier, the utilization of heuristics appears to be problem-dependent, and the study should be extended to other problems. Second, the sample selected for this study (and for many heuristic process studies) consisted of "better" problem solvers, and the investigator feels that the future studies should include subjects who are both "average" and "below average" mathematical problem solvers. Finally, an extension to include subjects in both elementary-junior high school and college could provide evidence concerning heuristic development during the complete schooling period. Such a study could provide evidence as to whether heuristics appear at different stages of schooling.

References

Blake, R. The effect of problem context on the problem solving ability of field independent-dependent students: A clinical study. Unpublished doctoral dissertation, The University of British Columbia, 1976.

Bloom, B. S., & Broder, L. J. Problem-solving processes of college students. *Supplementary Educational Monographs,* Chicago: University of Chicago Press, 1950, No. 73.

Gagné, R. M. Human problem solving: Internal and external events. In B. Kleinmuntz (Ed.), *Problem solving: Research, method, and theory.* New York: Wiley, 1966.

Garry, R., & Kingsley, H. L. *The nature and conditions of learning* (3rd ed.). New Jersey: Prentice-Hall, 1970.

Heathers, G. *Four process goals of education.* Center for Research and Development on Educational Differences, Harvard Graduate School, 1965.

Kantowski, M. G. Processes involved in mathematical problem solving. *Journal for Research in Mathematics Education,* 1977, *5*, 163-180.

Kilpatrick, J. Analyzing the solution of word problems in mathematics: An exploratory study. (Doctoral dissertation, Stanford University.) Ann Arbor, Michigan: University Microfilms, 1967. No. 68-06442.

Lucas, J. F. An exploratory study of the diagnostic teaching of heuristic problem-solving strategies in calculus. (Doctoral dissertation, University of Wisconsin.) Ann Arbor, Michigan: University Microfilms, 1972. No. 72-15368.

Polya, G. *Mathematics and plausible reasoning, Volume I. Induction and analogy in mathematics.* New Jersey: Princeton University Press, 1954. (a)

Polya, G. *Mathematics and plausible reasoning, Volume II. Patterns of plausible inference.* New Jersey: Princeton University Press, 1954. (b)

Polya, G. *How to solve it.* New York: Doubleday Anchor Books, 1957.

Polya, G. *Mathematical discovery, Volume I.* New York: Wiley, 1962.

Polya, G. *Mathematical discovery, Volume II.* New York: Wiley, 1965.

Webb, N. An exploration of mathematical problem-solving processes. Paper presented at the annual meeting of the American Educational Research Association, Washington, D.C., March, 1975.

The Alberta Journal of Educational Research *Vol. XXV, No. 4, December, 1979*

LAURENCE WALKER
Mount Saint Vincent University

Newfoundland Dialect Interference in Fourth Grade Spelling*

In this study, the influence of selected phonological features of the dialect spoken along the northeast coast of Newfoundland on fourth grade spelling performance was investigated. A test based on words whose pronunciation was affected by these features was administered to approximately 400 subjects from the dialect-speaking area and to a similar number of subjects in Nova Scotia. Using a standardized spelling test score as a covariate, analyses of the scores of 100-subject samples showed a significant effect in favour of the control group. Analysis of composition samples revealed patterns of misspelling that confirmed the major finding that phonological features of dialect speech did influence spelling errors.

Chomsky and Hallé (1968) claimed that English orthography is a near-optimal system for representing the spoken language. Part of the claim is based on the assumed power of a single, conventional system of spelling to represent, uniquely and without prejudice, all the various dialects of English. For the skilled reader it is of tremendous convenience to have only one writing system throughout the English-speaking world. Besides there is no convincing evidence, in spite of a large body of research, that for the beginning reader who speaks a nonstandard dialect, a written code which does not conform closely to his own dialect speech is a significant handicap. Less attention, however, has been paid to the question of whether nonstandard dialect speech interferes with the acquisition of the ability to spell conventionally. In the case of the beginning reader, recognition of a word involves going from a defined graphic stimulus to one of a range of permitted oral responses; for example, the written word "tomato" can be

---

Dr. Walker taught in public schools in Britain, Alberta, and Newfoundland before completing a doctorate at the University of Alberta in reading education. After three years as an assistant professor at Memorial University, Dr. Walker is now associate professor and Chairman of the Education Department at Mount Saint Vincent University, Halifax, Nova Scotia. The effect of dialect on spelling is part of his general research interest in children's use of written language.

* This study was funded by a grant from the Canada Council.

pronounced either with the American/tomeyto/ or the British/tomahto/. In other words dialect responses to printed words are psychologically valid if not always pedagogically so. For the naive speller, on the other hand, the task is one of going from one of a range of pronunciations to a single printed form. On the face of it, therefore, the nonstandard dialect speaker might be assumed to have more difficulty with spelling than with reading.

However, as teachers and spelling authorities have long recognized, the relationships between English sounds and spelling are not simple even for the speaker of the most conventional dialect. All spellers are faced with the difficulty of deciding which of several alternative representations of a particular sound or sequence of sounds is the conventional spelling. For the nonstandard dialect speaker the question is whether there is an additional difficulty in his making this choice owing to further divergencies between his dialect pronunciation and that of more standard dialects. The term "nonstandard dialect" is used to refer to a regional or social class variety of spoken English which has a number of distinguishing phonological, lexical, and syntactic features that set it off from other dialects. It is recognized that there is no single form of standard spoken English, and that so-called standard English, the variety of English spoken by the educated and influential members of a community, varies, mainly phonologically, from one geographical region to another.

The Role of Speech Sounds in Spelling

The larger question which underlies the issue of dialect interference in spelling is the extent to which the speller utilizes the sound information in the spoken word to be spelled. Traditionally, English was considered to be too irregular in its sound-letter correspondences for pronunciation to be a very useful aid to spelling. From an orthographic point of view, this perception was challenged by computer studies of large sets of words (Hanna, Hanna, Hodges, & Rudorf, 1966; Venezky, 1967). These studies revealed that when variables such as morpheme boundary, form class, stress, and graphemic environment are taken into account, English orthography demonstrates considerable regularity of sound-spelling correspondence. Venezky's (1967) conclusion that a morphophonemic level intervenes between the speech sound and its graphic representation was supported by a similar theory proposed by Chomsky and Hallé and applied to examples of derived word spellings by C. Chomsky (1970). This theory proposed that the phonological system includes an abstract lexical level from which surface phonetic pronunciations are generated, and that the spelling system represents this abstract level in a relatively consistent manner rather than the phonetic level in an irregular manner. For example, in "extreme" and "extremity" the second "e" represents two different phonemes; yet the spelling preserves the lexical relationship between the two words.

In a recent review of the issue of sound-spelling correspondences, Cahen, Craun and Johnson (1971) concluded that "overall the evidence indicates that phoneme-grapheme correspondences of some consistency do exist in English" (p. 297).

Spelling Processes and Dialect Interference

Having established that the English spelling system is related to sound, the question then becomes one of whether the speller makes use of the potential information in this relationship in order to help spell words. The extent to which he does and the manner in which he does is directly relevant to the question of dialect interference in spelling.

Theoretical analyses of the spelling process have been proposed by Simon and Simon (1973) and Peters (1970). Simon and Simon (1973) acknowledged that most people probably spell most words that are highly familiar to them by a direct recall method. Otherwise, they proposed, phonemic information is used in two ways: through stored phoneme-grapheme associations supplemented by rules governing the selection of different alternatives, and through a generate-test method analogous to the sounding out process in word recognition in reading (p. 119). This last strategy involves generating a list of possible spellings for an unfamiliar word and then using a visual recognition process, derived from experience with reading words, to select the appropriate form. The authors reported the use of a computer algorithm programmed to use both possible graphemic representations of phonemes and word recognition information to spell a body of words. The misspellings resulting from this computer program closely resembled those made by a group of grade 4 students on words that they had not previously studied. Simon and Simon (1973) concluded that this finding provided empirical support for the generate-test hypothesis as an explanation of "much of the spelling behaviour of fourth graders" (p. 135).

Peters (1970) proposed a similar procedure in which one of several possible orthographic representations is selected by the speller. However, unlike Simon and Simon, she based her selection procedure on serial probability information, or the awareness on the part of the speller of the most likely letter sequences permitted in the orthography. For example, "lodj" as a phonemic spelling contains the sequence "dj" which has a low or even zero probability in English orthography. Therefore spellers would likely reject it in favour of "lodge" whose "dge" sequence has a relatively high probability of occurrence. On the basis of the results of a letter sequence generalization task, included in a larger study of predictors of spelling achievement in 9-10 year olds, Peters (1970) claimed support for her hypothesis. However, no validity or reliability information on the task was offered in the report. Wallach (1963), earlier testing a similar hypothesis, concluded that the results of his study of fifth-grade spelling of different patterns of nonsense words were consistent with the proposition that good spellers as opposed to poor spellers are "learning a general property of English words, namely the transitional probability governing the sequential arrangement of letters in . . . words" (p. 61).

Thus, in approaching the question of dialect interference in spelling, the implications of three theoretical positions can be considered: the theory, more orthographic than psychological, that spelling represents an abstract lexical level in the phonological system; the generate-test theory; and the serial probability theory. Since the abstract lexical level theory is not addressed by this study, its implication for direct interference in spelling

will not be discussed here. Such a discussion is available in Brengelman (1970).

The generate-test theory of the spelling process would seem to have strong implications for the effects of dialect on spelling. In the stage where possible spellings are being generated, the set of possibilities produced by a dialect speaker for certain words might be expected to be different from the set produced by a standard-English speaker. For example, a speaker for whom a reduction rule deletes /d/ in word final consonant clusters might generate a list for / hænd / that would include "han." However, at the test stage "han" would likely be rejected since it would not be a word ever encountered in a reading situation; and, provided "hand" was among the options generated, the correct spelling would likely be selected. This theory would suggest that dialect speakers are likely to display spelling difficulties when their pronunciation of a word produces another word in standard English. For example, a speaker who pronounces "pig" like "peg" or "born" like "barn" may generate among his lists of alternative spellings forms that, while recognizable as familiar words, do not conventionally represent the appropriate lexical targets. Based on the generate-test theory, a dialect-based spelling error would require two conditions: first that the word to be spelled was not a familiar spelling word (in the case where it was, the word would presumably be spelled directly) and, second, that the word written as a misspelling was familiar in a reading or word recognition sense.

The serial probability theory would predict a wider range of dialect misspellings since the test stage in this view of the spelling processes relies on the likelihood of a sequence of letters being a permissible pattern in English. In this case "han" as a spelling for / hænd / would be predictable since the letter sequence is a common one.

Research into the influence of dialect on spelling performance has generally been concerned with testing the hypothesis that dialect does interfere rather than testing theoretical explanations of how any interference might occur. Several studies have provided evidence that some interference exists.

Graham and Rudorf (1970) studied grade 6 students in three geographical areas of the United States and found that dialect speech did appear to influence spelling. Wolfram and Whiteman (1971) found examples of misspelled words in compositions written by tenth-grade Black students that appeared to be related to dialect speech. Kligman, Cronnell, and Verna (1972) compared Black English-speaking grade 2 pupils with a sample of standard English-speaking pupils. When spelling errors on a multiple-choice test were divided into dialect-based and nondialect-based misspellings, they found a significant interaction between speaker group and error type. Nineteen per cent of the Black "dialect" speakers' errors were dialect-based as opposed to 12 per cent of the standard English speakers'. The study also revealed that certain dialect features were more prone to be implicated in misspellings than others. In a follow-up study, Groff (1978) found that by grades 4, 5, and 6, significantly fewer errors on the same test were dialect-related and that only a few of the dialect features seemed to be causing interference with spelling by the time pupils reached the middle grades. The test items which continued to produce dialect-related misspellings were:

"still," "walks," "streets," and "nests." It is interesting to note that the misspellings of these words which would be considered to be dialect-based would be respectively: "steel" or "steal," "walk," "street," and "nest." The fact that real word misspellings were the most persistent errors would be congruent with the generate-test theory proposed by Simon and Simon (1973). However, other test items such as "your" (you) and "its" (it) did not apparently continue to produce errors based on pronunciation.

In the case of other nonstandard dialects, a study by Boiarsky (1969) found that features of an Appalachian dialect were related to misspellings significantly more frequently in the spelling of tenth-grade students from rural Appalachia than in the spelling of students in a Philadelphia sample. An informal study in Newfoundland by Walker and Paddock (1975) found that of 253 spelling errors identified in stories written by a sample of 57 pupils from grades 3 to 6 in two rural schools, 62 errors (almost 25 per cent) appeared to be related to features of the local dialect.

In summary it seems fair to say that consistent but complex phoneme-grapheme relationships exist in English orthography. Different theories of the spelling process have suggested ways in which the speller learns to use the information provided by these relationships and the hypothesis could be derived from these theories that the speaker of a nonstandard dialect would have additional difficulties in learning to spell conventionally. Moreover, there is some evidence from studies of Black English speakers and of Appalachian dialect speakers that spelling errors are related to certain characteristics of divergent speech patterns. The present study sought to examine the question of interference in spelling as it applies to another dialect of English, that spoken in the Canadian province of Newfoundland. (The conditions under which the dialect developed are briefly described elsewhere [Walker, 1975].)

The Study

The purpose of the present study was to test the hypothesis that certain phonological features of a Newfoundland dialect cause spelling interference at the grade 4 level.

Dialect Spelling Test

The major source of data was a test of dialect spelling. Ten phonological features of the dialect spoken on the northeast coast of Newfoundland were selected (Paddock, 1977; personal communication, 1975). (Table 4 lists these features using Trager and Smith, 1951, notations.)

Word lists from three spelling series used in Newfoundland and the Province of Nova Scotia were then examined. All words whose pronunciation could be affected by one of the features shown in Table 4 and that were included in the grade 2-4 levels in these three series were listed. A random sample of the words was taken to produce a set of 36 test items with each of the 10 categories of features being represented in roughly the proportion revealed by the total population of words on the lists.

The 36 words were then put into random order and a context sentence written for each one. The words and sentences were tape recorded as an oral spelling test. A female, standard Canadian English-speaker did the

recording on good-quality tape in a university audiovisual studio. Each test word was pronounced, followed by the context sentence and the pronunciation of the word a second time. Sufficient time was allowed between the items for students to write the word down. As the result of a pilot test in one rural grade 4 classroom, two nondiscriminatory test items were replaced by other words from the same categories in the original lists. (The 36 test items are shown in Table 4.)

Population and Sample

The dialect-spelling test was administered to a sample of Newfoundland and Nova Scotia grade 4 students. The target population consisted of grade 4 students in three school districts on the northeast coast of Newfoundland. A control sample was drawn from the grade 4 students in two jurisdictions in Nova Scotia. Both areas contained rural populations engaged primarily in the fishing and lumbering industries. In both areas all schools containing grade 4 students, adjacent to salt water, and connected to the mainland by a road were identified. Table 1 shows details of the populations and samples in the two provinces. The Newfoundland schools were reduced by stratified random selection to a sample of 13 which reflected, in terms of size of school, the larger qualifying population of schools. Sixteen of the 17 Nova Scotia schools were retained in the control sample; the seventeenth, a small isolated school, could not be fitted into the data collection schedule. Table 1 also shows the precentage of grade 4 students in three types of schools: large schools, defined as those having more than 40 fourth graders; medium-sized schools with between 20 and 40; and small schools with fewer than 20. Also shown is the percentage of fourth grade students in multigrade classrooms. The figures show that the Nova Scotia sample contained higher percentages of students in medium-sized and small schools and in multigrade classrooms than the Newfoundland sample. Another source of difference was the fact that, although the economies of the two regions were based on the same

TABLE 1
POPULATION AND SAMPLE

	Newfoundland	Nova Scotia
Population		
Schools	28	17
Grade 4 students	900	450
Sample		
Schools	13	16
Grade 4 students tested	391	430
Final Sample	100	100
Percentage of students in:		
large schools	62	38
medium schools	26	45
small schools	12	17
multigrade classes	26	34

industries, the Nova Scotia control sample was drawn from a more prosperous population than the Newfoundland sample.

Although complete data were collected from 391 Newfoundland students and 430 Nova Scotia students, analysis was carried out on randomly-selected 100 students from each sample.

Data Collection

Apart from the dialect-spelling test, two other sets of data were collected. To obtain a measure of general spelling performance, the Level 2, Form A Spelling Sub-test of the *California Achievement Tests* (Tiegs & Clark, 1970) was administered. This test requires students to identify a misspelled word from sets of 4 words. None of the misspelled words in the test appeared to have any coincidental connection with the phonological features used to derive the dialect-spelling test.

As a source of data on dialect intrusion into spelling under ordinary writing conditions, a sample of free composition was collected from each subject. A writing task was designed and followed carefully in each classroom. Some soap bubbles were blown and the students allowed to watch, touch, smell, and taste them. Then a discussion was encouraged followed by a short story about a bubble, read aloud to the class. At this point the students were asked to make up and write their own stories about a bubble. No help was given with spelling; students who asked for help were told to do their best. No time limit was imposed on the writing.

All data were collected within a ten-day period in April, 1976. Collection was carried out by two research assistants, one in Newfoundland and the other in Nova Scotia. Both were females with teaching qualifications and experience. Data were collected in two sessions with each class.

Data Analysis

Analysis of covariance was used to test for significant differences between the two samples on the dialect-spelling test. Scores on the standardized spelling test were used as the covariate.

All misspellings on the free compositions were tallied and those which appeared to be related to any of the 10 phonological dialect features involved in the study were noted. The proportions of misspelled words that could be attributed to these dialect features were compared using a *t* test.

Results

Dialect-Spelling Test Scores

The results of the analysis of covariance of the scores on this 36-item test are shown in Table 2. The means and standard deviations for dialect-spelling scores are shown in Table 3.

Table 2 shows that there was a significant difference between the two samples on spelling ability as measured by a standardized test. The mean score for the Newfoundland sample was 16.46 with a standard deviation of 5.37; for the Nova Scotia sample the mean was 18.21 with a standard deviation of 4.14. Clearly, therefore, the samples were unequal with respect to general spelling performance, making it necessary to control for spelling

TABLE 2
ANALYSIS OF COVARIANCE: DIALECT-SPELLING TEST SCORES BY PROVINCE WITH STANDARDIZED SPELLING TEST SCORES AS A COVARIATE

Source	SS	df	MS	*F*	*p*
Covariate	7067.26	1	7067.26	252.86	.001
Province	875.74	1	875.74	31.33	.001
Residual	5506.02	197	27.95		

TABLE 3
DIALECT SPELLING TEST SCORES: MEANS AND STANDARD DEVIATIONS

Province	*N*	Mean	SD
Newfoundland	100	24.44	9.24
Nova Scotia	100	30.70	5.54

ability. The Pearson *r* correlation for the total combined sample (n = 200) between the standardized spelling scores and the scores on the dialect spelling test was .72.

The main finding shown in Table 2 is that, over and above general spelling ability, there was a significant difference between the two groups on the dialect-spelling test. Table 3 shows that this difference favoured the control sample. In other words, on a spelling test made up of words deliberately chosen because logically they would be susceptible to misspelling by Newfoundland dialect speakers, the Newfoundland sample performed significantly lower than the Nova Scotia sample even when differences in general spelling ability were controlled. The design of the study suggests that this finding can be explained by the dialect differences between the two samples and that the phonological features of the nonstandard Newfoundland dialect did contribute to the spelling difficulties of the Newfoundland fourth graders.

Table 4 shows details of the performance of each group on the 36 dialect-spelling test items and overall on the 10 feature categories. Eleven of the test items produced a substantial number of misspellings that were predictable from the dialect features involved: "poor," "pool," "bill," "still," "kid," "fill," "thick," "both," "stem," "picked," and "spelled." Six of the ten phonological categories were responsible for substantial percentages of dialect-based misspellings. However, in the case of the one-word category /uw/ → $\left\{ \begin{matrix} /ɔ/ \\ /ʌ/ \end{matrix} \right\}$ /——r the control group produced a higher percentage of errors that were apparently dialect-based than the target group, albeit with a smaller number of errors being involved. In the other five sensitive categories the control group's

TABLE 4
TOTAL SPELLING ERRORS AND DIALECT SPELLING ERRORS BY PHONOLOGICAL CATEGORY

Category/word	Newfoundland (n = 100)		Nova Scotia (n = 100)	
	Total Misspellings	Dialect Misspellings	Total Misspellings	Dialect Misspellings
1. /ɔ/ → /æ/ / ___r				
corn (carn)	17	2	5	0
born (barn)	26	0	7	0
morning (marning)	35	0	16	0
corner (carner)	33	0	25	0
Category percentage		1.8%		0%
2. /uw/ → {/ɔ/, /ʌ/} / —r				
poor (pore, pur)	24	13	11	8
Category percentage		54.2%		72.7%
3. /iy/→/e/ / — r				
deer (dare)	23	1	13	1
Category percentage		4.3%		7.7%
4. /uw/→{/u/, /o/} / —l				
pool (pull, poll)	24	17	6	2
cool (cull, coll)	48	5	12	3
Category percentage		30.6%		27.8%
5. /ɪ/ →{/iy/, /e/} / —l				
bill (beel, bell)	42	26	3	2
still (steel, stell)	38	21	8	1
kill (keel, kell)	42	12	0	0
fill (feel, fell)	36	24	11	1
Category percentage		52.5%		18.2%
6. /θ/ → /t/				
thick (tick)	29	12	9	1
everything (everyting)	57	3	31	1
both (boat)	39	8	22	5
teeth (teat)	23	4	11	1
Category percentage		18.2%		10.9%
7. /ð/ → /d/				
these (dese)	45	0	19	0
grandmother (grandmodder)	30	1	13	0
Category percentage.		1.3%		0%
8. /h/ ↔ zero				
evening (hevening)	24	0	20	0
afternoon (hafternoon)	25	0	7	0
ice (hice)	2	0	1	0
hay (ay)	28	1	4	0
hang (ang)	27	1	12	0
hole (ole)	51	2	13	0
Category percentage		2.5%		0%
9. /e/ → /ɪ/ .				
led (lid)	48	0	60	1
pet (pit)	4	1	1	0
letter (litter)	27	4	11	2
yesterday (yisterday)	34	0	14	0
stem (stim)	46	14	23	4
shell (shill)	21	4	10	0
Category percentage		12.8%		5.9%
10. consonant cluster reduction				
picked (pick)	47	32	25	14
spelled (spell)	42	15	14	7
sold (sole)	13	1	16	1
pound (poun)	41	5	21	0
salt (sal)	31	0	22	1
coast (cose)	34	2	31	0
Category percentage		17.9%		17.8%

percentage of dialect-based spelling errors was lower than the Newfoundland sample's and involved considerably fewer misspellings. Within the consonant cluster reduction category, the greatest effect seemed to involve the -ed form of the verb. There was little or no evidence that the free alternations between / ɔr / and / ær /, / iyr / and / er /, and / ð / and / d /, and the difference in the Newfoundland use of initial / h / affected spelling performance.

TABLE 5
SPELLING IN FREE COMPOSITION

	Newfoundland (n = 100)		Nova Scotia (n = 100)	
	Mean	S.D.	Mean	S.D.
Composition length (words)	78.3	44.6	114.2	50.8
Misspelled words	8.3	4.6	7.3	5.8
Dialect-related misspelled words	1.12	1.64	.39	.72
Percentage of misspelled words that were dialect-related	12		5	

Free Writing Misspellings

Table 5 shows the results of the descriptive analysis of the free compositions written by the two samples. The Newfoundland compositions were on the average shorter in length; they contained more misspelled words; and a larger percentage of their misspelled words could be related to the phonological features of the dialect. A t test of the difference between the two samples on the percentage of spelling errors attributable to dialect was significant beyond the .001 level (t = 3.76, df = 198). These findings indicate that the Newfoundland students made spelling errors in their free compositions that were significantly more likely to be relatable to phonological features of their dialect speech than the spelling errors of the control sample.

Discussion

Clearly the Newfoundland sample experienced more spelling difficulties with the dialect-spelling test than did the Nova Scotia sample, even when spelling ability, as measured by a standardized spelling test, was statistically controlled. Since the items on the dialect-spelling test were selected on the basis of their relationships to certain phonological features of the dialect speech of the Newfoundland sample, it seems logical to assume that the dialect speech was causally related to the spelling difficulty. This finding of dialect interference is congruent with those of studies with other dialects (Kligman et al., 1972; Groff, 1978; and Boiarsky, 1969). Moreover, the finding that a significantly greater percentage of spelling errors made by the Newfoundland fourth graders in a free writing exercise were dialect-based indicates that the interference is not limited to spelling

isolated words in a controlled exercise but that it represents an extra impediment to these students' writing.

Not all the dialect features produced evidence that they were related, at least in a direct way, to spelling difficulty. That is not to deny, however, that the features which appeared not to produce dialect-based errors could have exercised an indirect influence on learning to spell. For example, one could speculate that a dialect speaker, faced with additional instances where the orthography does not consistently represent his speech sounds, could suffer a loss of confidence in the orthographic system such that his mastery in both reading and writing is delayed. In this case misspellings of a more random and unpredictable nature could well result from a general sense of dismay at the task of writing words. It was certainly true that the dialect speakers in this study, the Newfoundland fourth graders, were considerably less adept at spelling than their Nova Scotia counterparts.

The finding that only some features of the dialect interfered with spelling was similar to the results of studies with Black English (Kligman et al., 1972; Groff, 1978). Both of these studies found that the / ɪ / sound preceding / l / as in / stɪll / produced "steal" and "steel" misspellings among both the Black English-speaking second graders and middle school students studied. In the present study seven out of 38 misspellings of the word "still" were of this nature. However, in the case of this Newfoundland dialect the / ɪ / may go either /iy/ or /e/ when followed by /l/. This extra variation showed up as 14 misspellings of "still" as "stell" in the Newfoundland sample. The same pattern was shown by the Newfoundland misspelling of the other words illustrating this feature: "kill" produced one "keill" and nine variations of "kell"; "fill" was misspelled as "feil" once and as "fell" 20 times; and "bill" appeared once as "beil" and 25 times as "bell." Another feature of Black English which Groff (1978) found persisting in the spelling errors of his middle grade sample was the / d / and / t / deletions at the end of past tense forms of verbs. In the present study two items involving this feature produced the largest number of dialect-based misspellings. To a reduced extent the same misspellings were present in the errors of the control group. Kligman et al. (1972) concluded that the / e / sound before nasals as in /ten/ caused difficulty for Black English-speaking second graders and some difficulty for standard English-speaking speakers. However, Groff (1978) found that by the middle grades this difficulty had largely been overcome by his sample. The word "stem" was an item in this study that produced a relatively large number of / e / → / ɪ / misspellings among the Newfoundland sample and a small number among the control group. The same word feature in other phonological environments resulted in only a few predictable errors. The similarity between these selective interference findings may indicate that some features in some environments are intrinsically difficult to perceive and that spelling problems result not simply from the fact of dialect variation but from this variation combined with other factors of difficulty.

Some details of the results of this study may be discussed in terms of theories of the spelling process as they apply to dialect interference. The generate-test theory would predict, in cases where words were unfamiliar as spelling words yet familiar as reading words, that dialect-related mis-

spellings would be more frequent when the error is itself an alternative word. The serial probability theory would, on the other hand, predict errors that are congruent with permissible letter sequences in English. An informal analysis of the 231 dialect-based misspellings produced by the Newfoundland sample showed that 65 per cent of the responses were real words and 35 per cent were nonwords. That is, the majority of errors that the dialect would predict were responses like "bell" for "bill" and "tick" for "thick" as opposed to nonword responses like "beal" and "tek." However, there were some individual words that did not conform to this pattern: "still" was more often spelled as "stell" than "steel" or "steal," and "stem" was realized as "stim" by 14 subjects. Thus nothing more than very tenuous support for the generate-test theory is offered by these incidental findings. Further research might find it useful to pursue the issue of the relative power of the generate-test and the serial probability theories to explain dialect interference in spelling.

The results of this study have implications for the general issue of dialect interference in spelling and for spelling instruction in Newfoundland. From the consistent findings over at least three dialects, the fact of dialect interference in spelling seems well established. However, this is only a rather crude answer to a complex question as the variation among dialect features in their effect on spelling suggests. This variation seems to be both between phonological features and within particular features as inconsistent results between words illustrating the same features are revealed. The research question now should be what features interfere in what circumstances or environments and, from a theoretical perspective, how does this interference occur?

Dialect interference in spelling would seem to be another reason why teachers of language arts should know the features of the local dialect. On Newfoundland's northeast coast, teachers who have a systematic knowledge of the dialect can adopt a more understanding attitude to spelling problems involving past tenses of verbs, front vowel sounds, and consonant cluster reduction. Such teachers would be able to scrutinize spelling texts and workbooks to identify words that contain dialect pitfalls not anticipated by authors from central Canada or the United States. They would be able to adjust their instruction to help students around the friction points between their dialects and more conventional pronunciation as they affect spelling. The necessary knowledge would call for inservice as well as preservice education in dialectology and its implications for language learning, especially for learning to use written language.

References

Boiarsky, C. Consistency of spelling and pronunciation deviations of Appalachian students. *Modern Language Journal,* 1969, *53* (5), 1113-1118.

Brengelman, F. H. Dialect and the teaching of spelling. *Research in the Teaching of English,* 1970, *4*, 129-138.

Cahen, L. S., Craun, M. J., & Johnson, S. K. Spelling difficulty: A survey of the research. *Review of Educational Research,* 1971, *41* (4), 281-301.

Chomsky, C. Reading, writing, and phonology. *Harvard Educational Review,* 1970, *40* (2), 287-309.

Chomsky, N., & Hallé, M. *The sound pattern of English*. New York: Harper and Row, 1968.

Graham, R. T., & Rudorf, E. H. Dialect and spelling. *Elementary English,* 1970, *47,* 363-376.

Groff, P. Children's spelling of features of Black English. *Research in the Teaching of English,* 1978, *12* (1), 21-28.

Hanna, P. R., Hanna, J. S., Hodges, R. E., & Rudorf, E. H. Phoneme-grapheme correspondences as cues to spelling improvement. Washington: U.S. Government Printing Office, 1966.

Kligman, D. S., Cronnell, B. A., & Verna, G. B. Black English pronunciation and spelling performance. *Elementary English,* 1972, *49* (8), 1247-53.

Paddock, H. Preliminary dialect mapping of Newfoundland. In H. Paddock (Ed.), *Languages in Newfoundland and Labrador.* (Preliminary Version.) St. John's: Memorial University Department of Linguistics, 1977, 90-99.

Peters, M. L. Success in spelling: Study of factors affecting improvement of spelling in the junior school. *Cambridge Monographs on Education No. 4.* Cambridge: Cambridge Institute of Education, 1970.

Simon, D. P., & Simon, H. A. Alternative uses of phonemic information in spelling. *Review of Educational Research,* 1973, *43* (1), 115-137.

Tiegs, W., & Clark, W. W. *California Achievement Tests, 1970 Edition.* Monterey, California: CTB/McGraw-Hill, Inc., 1970.

Trager, G. L., & Smith, H. L., Jr. *An outline of English structure.* Norman, Okla.: Battenburg Press, 1951.

Venezky, R. L. English orthography: Its graphical structure and its relation to sound. *Reading Research Quarterly,* Spring, 1967, *2* (3), 75-105.

Walker, L. Newfoundland dialect interference in oral reading. *Journal of Reading Behavior,* 1975, *7* (1), 61-78.

Walker, L., & Paddock, H. Spelling and the Newfoundland dialect. *Morning Watch,* 1975, *2* (4), 9-13.

Wallach, M. A. Perceptual recognition of approximations to English in relation to spelling achievement. *Journal of Educational Psychology,* 1963, *54* (1), 57-62.

Wolfram, W., & Whiteman, M. The role of dialect interference in composition. ERIC Document, ED 045 971, 1971.

T. E. KIEREN
The University of Alberta
and
B. SOUTHWELL
Nepean College of Advanced Education

The Development in Children and Adolescents of the Construct of Rational Numbers as Operators

The purpose of this study was to characterize the growth of the operator construct of rational numbers in children and adolescents. A random sample of 72 subjects from 4 elementary and junior high schools were interviewed while doing 7 direct and 7 inverse tasks in either a "real problem packing machine" or a number pattern setting. There were significant age group performance differences. There was an important age × representation interaction effect. Children in the machine setting achieved relatively high scores at age 11-12 which were not duplicated by the pattern group until age 13-14. A nominal analysis revealed 3 phases of rational number performance: a "½" phase, a unit operator phase, and a general operator phase. It appears that the machine setting prompted a functional interpretation of equivalence and allowed for successful use of covaried partitioning—these two factors promoting earlier general operator performance. The pattern setting appeared to prompt a more algebraic interpretation of equivalence which was used very successfully by older students, particularly females.

Fractional or rational number ideas historically have been and continue to be difficult for students in school mathematics. Yet, rational numbers are critically important because they represent a means by which a person can deal numerically with continuous phenomena in the same way counting number ideas can be applied to discrete phenomena.

Dr. Kieren completed his doctoral degree at the University of Minnesota. He has been associated with the Department of Secondary Education at the University of Alberta for most of the past 12 years. His research has focussed on the mathematical thinking of children and young adults.

Dr. Beth Southwell completed her doctoral degree at the Institute of Education, University of London, and is currently a Senior Lecturer at Nepean College of Advanced Education, Australia. Her research has centred around the development of mathematical concepts in children and includes investigations in Papua New Guinea.

One reason that the rational number construct proves difficult to learn is its complexity. Kieren (in press) has theoretically posited the existence of five subconstructs which appear necessary for a fully functional fractional number construct. These five are part-whole, ratio, quotient, measure, and operator. Full development of these five constructs by a person should allow him or her to apply these ideas in the wide variety of daily and mathematical circumstances where they are pertinent. Kieren (in press) has also hypothesized the existence of mechanisms which a person uses to build up these constructs. Some of these mechanisms, for example, proportional reasoning, are *developmental* in nature and depend on the onset of various stages (to a certain extent age-related) for their personal usefulness. Noelting (1978a) has carefully documented the characteristics of the stages of such reasoning as it pertains to fractions. Kieren (in press) labels other mechanisms as *constructive*. He theorizes the existence of mechanisms for developing rational number constructs which would parallel the central role played by counting in whole number development.

The study reported here focusses on the development of the operator subconstruct of rational numbers. It represents an extension of a previous study of rational numbers as operators (Kieren & Nelson, 1978) in that it attempts to study more clearly and carefully phases in development and mechanisms involved.

The operator construct is an important focus of study for two reasons. First fractional or rational numbers are by nature two-sided; they are simultaneously additive vectors and functions which can be composed. The operator subconstruct is the clearest representation of rational numbers as numbers to be multiplicatively composed. The second reason for their importance is curricular. Approaching the study of fractions through operators has been suggested by Dienes (1971) and widely implemented in the school curriculums of the Federal Republic of Germany (Padberg, 1978). Like many other approaches to the teaching of fractions, it is important to see how and why it might be made most effective in terms of the learner's construct development.

There are two major purposes of this study. The first is to trace the development of the operator construct. Does the quality of this construct change with age and, in particular, are there quantitatively identifiable growth points or levels as tentatively seen in a previous study (Kieren & Nelson, 1978)? How do developmental and constructive mechanisms, if observable, affect this growth?

The second major purpose deals with the representation of the construct. In Kieren and Nelson (1978), the fractional operator was represented by a machine, the nature of which the young person was to discover. This representation allowed the person to manipulate objects and see the idea in a "real" problem setting. Such settings prove difficult to include in curricular materials. A simpler curricular approach to operators is found in the study of patterns in pairs of numbers. An important question is, "How does the pattern representation as opposed to the machine representation affect the way in which the young person thinks about the fractional number?"

There are three minor purposes of the study. In a previous study Kieren

and Nelson (1978) found that the notion of "1/2" formed a stage or level in fractional thinking in the operator setting. This is confirmed by the work of Piaget, Inhelder and Syzminska (1966) and Noelting (1978a) in other ways. Kieren and Nelson further observed that older children and adolescents persisted in making "1/2 errors" in situations they found unfamiliar. Because in the previous study the first task was a "1/2 task" it was thought that possibly this task induced a "set" which crept into later tasks. In the current study, subjects were randomly assigned to one of two task orderings, one which started with a "1/2 task" and one which did not.

A second minor purpose was to further establish the fact that a subject who could handle the direct fractional operator could also do its inverse (Kieren & Nelson, 1978). This result is seen as a particular manifestation of the mechanism of reversibility and should be well established in the subjects of ages used in this study.

The third minor purpose was to look at the notion of composition of functions. In other mathematics educational research (Lovell, 1971; Thomas, 1975) composition of function is seen as an advanced notion. This present study allowed one to consider composition in the setting of fractional operators and is indeed central to this construct. (A detailed report of this aspect of the study can be found in Kieren and Southwell, 1978).

Procedures

Sample

As in the Kieren and Nelson (1978) study, the sample for this study was chosen from students in Grades 4 to 8. This was done because these are the levels of most significant instruction in rational numbers. In addition, students at these ages should have reached the developmental level in which they can exhibit reversibility and have primitive proportional reasoning skills (Noelting, 1978a). Thus, 72 subjects were selected randomly from the populations of two elementary and two junior high schools. There were 16 subjects at each of Grades 4, 5, and 6 and 12 subjects at Grades 7 and 8. One elementary and one junior high school were from the Public schools while the other two were from the Catholic school system of a large western Canadian city. The schools were in either inner city or suburban locations. For the purpose of analysis this sample of 72 was separated by age into five categories as listed in Table 1.

To test the appropriateness of the lower bound of the age range, four children from Grade 2, aged 7 and 8, also were tested. In addition, 4 high achieving Grade 10 students, aged 16, were tested. In the previous study there was only one person, aged 13, who reached the highest performance category. The four Grade 10 students were tested to see if they would perform uniformly at a high level and thus further validate the task set.

TABLE 1
AGE GROUP SIZES

Age in years	< 10	$10 \leq a < 11$	$11 \leq a < 12$	$12 \leq a < 13$	$a \geq 13$
n	11	20	13	13	15

The testing was done during April and May of 1978 in a manner documented below by the two experimenters using a well-defined protocol. Each child was tested separately in a room provided by each of the schools. All of the children in the sample had received some instruction on fractions or rational numbers and this was uniform, at least in terms of objectives, across schools at a grade level. In all cases, the subjects were drawn from more than one classroom in a school.

Upon coming to the testing area, students were randomly assigned to either a machine or pattern setting and, within settings, assigned to either a "1/2 first" or "1/3 first" task sequence. These procedures resulted in the categories outlined in Table 2. The two experimenters randomly administered the machine and pattern settings.

TABLE 2
NUMBERS OF SUBJECTS ASSIGNED TO TESTING SETTINGS

	Machine	Pattern
1/2 first	19	17
1/3 first	15	21

Instruments

Machine. A series of fractional tasks was presented using two different modes. The first mode involved the use of a coloured box with input and output doors. The subjects were told that the box represented a packing machine which was programmed to behave in a certain way. As objects, represented by small pieces of cardboard, are fed into the machine, they are "packed" into larger packages, represented by a larger piece of coloured cardboard.

The subjects were asked to feed in a certain number of objects and count the resulting output. After up to six trials of this kind, they were asked to predict what the outcome would be on five specific tasks presented to them on a sheet with IN and OUT columns. Upon completion of the sheet, each subject was asked to explain how the machine was working. A written record of free responses was kept.

A second sheet with IN and OUT columns was handed to each subject for completion. "What would I put into the machine to get 7 out?" was the form of question asked. Upon completion of the five tasks on this second sheet, the subject was again asked to explain how the machine was working and his free responses were recorded.

The subject was then told that the machine had been reprogrammed and a similar sequence of questions and instructions were given using a different fractional operator. In all, the subject was presented with 7 sets of 5 direct and 5 inverse items, representing a total of 14 tasks. Variables used in these tasks were (i) whether the operator involved was a unit (1/2, 1/3, 1/3 × 1/2) or a non unit fraction (3/4, 2/3, 3/4 × 1/2, 3/4 × 2/5) and (ii) whether the operator was single (1/2, 1/3, 2/3, 3/4) or composite (1/3 + 1/2, 3/4 + 1/2, 3/4 + 2/5). (See Kieren and Nelson, 1978, for details of this procedure.)

In the latter case, two machines were used. The first, with which the subjects were already familiar, put objects into packages. The output of this first machine became the input for the second machine which packed the packages into cartons (represented by larger pieces of cardboard). For instance, for the composition, $3/4 \times 1/2$, sixteen objects were fed into the first (3/4) machine and twelve packages came out. These twelve packages were fed into the second machine and six cartons resulted. As in the former setting with one machine, the subjects were given up to six trials and were then asked to predict the outcome in five direct and five inverse tasks on this composite setting.

A third variable was whether subjects were presented with a sequence of tasks beginning with "1/2" or whether they were presented with a sequence beginning with "1/3". In the former case, the sequence was 1/2, 2/3, 1/3, 3/4, $1/3 \times 1/2$, $3/4 \times 1/2$, $3/4 \times 2/5$. The second sequence was 1/3, 3/4, 1/2, 2/3, $1/3 \times 1/2$, $3/4 \times 1/2$, $3/4 \times 2/5$.

The sets of tasks were presented to the subjects individually. The subjects were given as much time as they needed to complete as many of the tasks as possible. Each subject followed one of the above sequences and proceeded until he or she was unable to make any predictions or explain how the machine(s) was functioning. Predictions were recorded by each subject on the appropriate sheets for the 14 (7 direct, 7 inverse) tasks. In addition, the experimenter kept a written record of subjects' discussion and descriptions of the machine.

Pattern. The second mode involved the use of patterns displayed on pieces of cardboard, with numerals linked together by an arrow, e.g.

$$12 \rightarrow 4$$

Up to six cards with the same pattern (e.g., 1/3) were displayed and the subject asked to find the pattern. This was done in the same way as with the machine mode. After several cards were displayed, the subject was asked to predict the outcome using the same pattern; for example

given $12 \rightarrow 4$,
$9 \rightarrow 3$,
$27 \rightarrow 9$,
$60 \rightarrow 20$,
and $120 \rightarrow 40$,
$6 \rightarrow ?$ etc.

The subject recorded his or her predictions on five tasks on sheets as with the machine mode. For inverse items, the subject was asked what would have to be on the left hand side of the arrow to give n on the right hand side.

The tasks presented to the subjects using the pattern mode were the same as those given using the machine. Thus, the same three variables were present in this setting. The subjects recorded their predictions individually and the experimenter kept a written record of discussion and descriptions used by the subjects. These test items were such that the required result was always a whole number. In addition, the items were such that three items tested the operator on "small" numbers ($n \leqslant 40$), one on numbers from 50 to 200, and one item on numbers in the thousands. This was done to assess whether the subject was using a single strategy for each operator and

whether this strategy was personally applicable across a wide domain of numbers. Thus for small numbers a subject might see a coincidental but successful pattern which would not carry over to larger numbers. Each subject was given as much time as was required to complete the tasks and proceeded until unable to go any further. In analysis of results, mastery was assessed on success of 4 of 5 items, at least one of which would have to incorporate a "large" number.

Both the machine and the pattern mode proved to be viable settings. All subjects tested were able to function within these settings. The pattern mode, however, is considered to present a more abstract task than the machine in which the subjects were able to manipulate concrete objects providing both motivation and clarification of the task. The time for each interview was dependent on the needs of the subject; actual times ranged from 13 to 70 minutes.

All students did not complete all settings. For example, subjects who were unable to do 3/4 tasks were not asked to attempt $3/4 \times 1/2$ or $3/4 \times 2/5$. The minimum number of tasks attempted was one and the median number, 8 of 14.

Questions and Hypotheses

In an important sense this research was hypothesis-generating. It allowed for a detailed consideration of child and adolescent behaviour against controlled protocols with two representations of the operator construct of fractional or rational numbers. From the records of subject behaviour, one can attempt to develop a reasoned explanation or set of hypotheses which describe the pattern of behaviour. The following hypotheses assisted in the development of such an explanation.

—There is no difference in mean performance by students faced with "1/2 task first" sequence as opposed to the "1/3 task first" sequence.

—Performance on inverse tasks is independent of performance on direct operator tasks.

—There exist differences in mean performances of groups at different age levels with older subjects performing at a higher level.

—There are no differences in mean performance by subjects in the machine and pattern settings.

—There are stable categories of performance which typify the development of the operator construct.

The data used in testing these hypotheses and in generating the theoretical discussion below come from two sources. The performance of subjects on each of 14 tasks (5 items per task) was recorded. The split halves reliability of results on these items is .91. The transcripts of subjects' verbal and nonverbal reactions to each task faced were also available for study.

Results

Preliminary

Before turning to the major purposes of the study, two preliminary results are of interest. The first deals with the testing sequence. Because "1/2" holds such an important position in rational number construct development,

the "set" generated by presenting the "1/2 task" first may have affected the results of the study. Thus the mean performances of the "1/2 first" vs. "1/3 first" were compared at each of the age levels of the study. With the exception of youngest age level ($a < 10$) where the "1/2 first" group was higher, (t = 2.77, sig. at .02) there were no significant differences in the mean performances of the two groups at any age level. The results in the youngest age group were greatly affected by the performances of one person in each sequence category. This and a study of 1/2 errors in subject protocols leads to the conclusion that, as will be seen below, although "1/2" is a stage in development, the sequence "1/2 first" does not account for differing performances or errors among students.

The second preliminary result deals with performance on direct and inverse items. Table 3 shows the number of subjects mastering (4 of 5 items correct) related direct and inverse tasks (for example, numbers of subjects mastering both "1/2 direct" and "1/2 inverse" items). Thus, although the inverse items presented somewhat more difficulty than direct items (42.5% vs. 48.5% mastery performances), there did exist a condition of dependency. Therefore a subject who mastered a direct task likely would also master the related inverse task. It appears that it is the nature of the operator (1/2, 3/4, etc.) which determines performance rather than the aspect of inverse within an operator.

TABLE 3
NUMBERS OF PERFORMANCES IN VARIOUS MASTERY CATEGORIES WITH RESPECT TO PAIRED DIRECT AND INVERSE TASKS

		Inverse: Mastered	Inverse: Non-mastered	
Direct: Mastered		216	42	258 (48.5%)
Direct: Non-mastered		10	264	274 (51.5%)
		226 (42.5%)	306 (57.5%)	532

$p(\chi_2 = 351) < .01$

A final preliminary result relates to the performance of Grade 2 students (ages 7, 3 to 8, 5). Only one of 4 students was able to do any of the tasks, a girl aged 7 years 3 months, and she was able to cope only minimally with the 1/2 task. This result is as hypothesized and is consistent with the findings of Noelting (1978a). He found that the onset of the first level of the equivalence class reasoning stage was at age 8 and the more general equivalence class stage at age 10. Thus the younger students would, at best, be able to cope with the "1/2" tasks of this study.

Central Results

After up to six trials with feedback on each of the task settings, the subject was asked to predict, in writing, the output (or input) of a machine or pattern for a given task. There were 5 direct and 5 inverse items per task.

TABLE 4
MEAN PERFORMANCES OF SUBJECTS IN VARIOUS AGE AND REPRESENTATIONAL CATEGORIES

Mode	Ages in Years					Total
	a < 10	10 ≤ a < 11	11 ≤ a < 12	12 ≤ a < 13	a ≥ 13	
Pattern	11.4	26.1	37.8	32.3	48.3	31.2
Machine	14	24.8	43.2	42.6	40.9	32.9
Total	12.6	25.5	40.7	36.3	44.4	

TABLE 5
ANALYSIS OF VARIANCE OF ACHIEVEMENT SCORES USING REPRESENTATION AND AGE AS BLOCKING VARIABLES

Source	Sum of Squares	df	Mean Square	*F*
Representations	213	1	213	.88
Age	7233	4	1808	7.47*
Interaction	3902	4	975	4.03*
Within Cells	15034	62	242	
Total	26382	71		

* Significant at a .01 level.

Thus there was a possible total of 70 correct predictions. Table 4 presents the mean achievement of each age level group as well as for each age level group in each of the representational settings.

Table 5 reports the analysis of variance performed using representation and age as variables.

It is clear from this analysis that there is a significant age effect as would be expected. While there appears to be no general representation effect, this result is tempered by the interaction which can be vividly seen in Figure 1. From this figure and from related contrast analyses, the following conclusion can be drawn. For the machine groups there are significant achievement gains from age (<10) group to age ($10 \leqslant a < 11$) group and from that group to the age ($11 \leqslant a < 12$) group, the latter a dramatic change. There is no significant growth seen at age ($12 \leqslant a < 13$) or age ($a \geqslant 13$) groups. With the pattern groups significant gains appear from age (<10) to age ($10 \leqslant a < 11$), from age ($10 \leqslant a < 11$) to age ($11 \leqslant a < 12$) and a dramatic gain from age ($12 \leqslant a < 13$) to age ($a \geqslant 13$). Thus, it appears that the machine group reaches a high level of performance at age 11-12 while with the pattern group, high level of performance comes at a later age ($a \geqslant 13$).

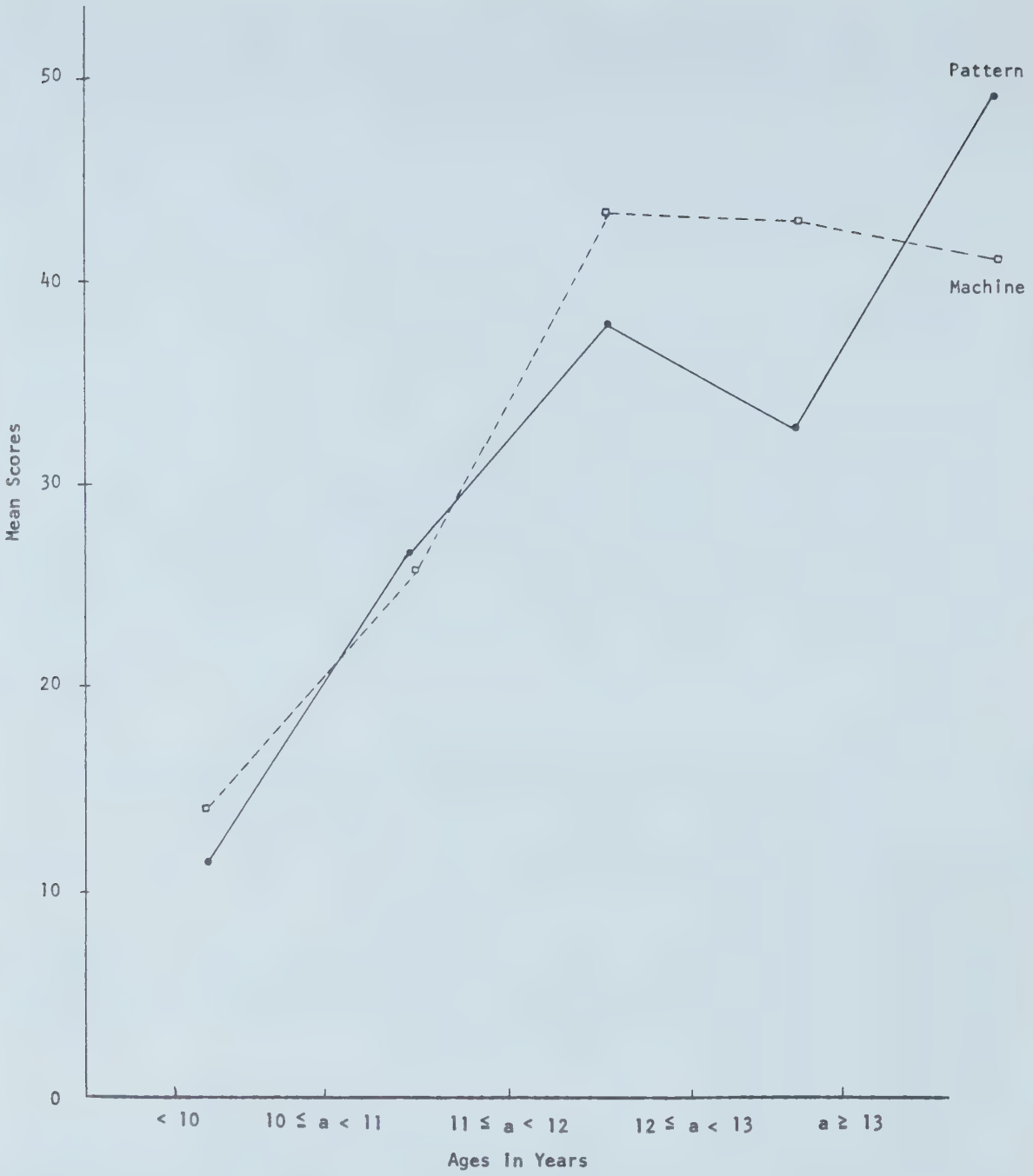

Figure 1. Performances by age and representation.

Nominal Analysis

The task performances and explanations given by children and young adults in each of the settings were analyzed to determine if certain "stages" in the development of the construct might be observed. From this analysis seven levels or stages appeared.

I: Mastered none of the 14 tasks.

II: Mastered the "1/2" tasks (direct or inverse) only.

III: Mastered "1/2" and "1/3" tasks.

IV: Mastered unit tasks including composition.

V: Mastered some non-unit tasks (2/3, 3/4 compositions).

VI: Mastered all tasks except 3/4 × 2/5.

VII: Mastered all tasks.

A purpose of this study was to search for performance categories. Hence no formal attachment of criterion tasks to stages was made a priori and the corresponding hierarchy analysis was not formally considered. It is interesting to note that well over 90% of subjects in level n performed all of the tasks associated with level $n - 1$. Conversely less than 10% of the subjects who failed a level n task mastered any task at a level above level n. For example all subjects who mastered 3/4 also mastered 1/2 × 1/3, and of those failing "1/2 × 1/3," only 1 subject mastered a non-unit task. Thus, though not formally tested, the levels appear to be hierarchical in nature.

TABLE 6
LEVELS IN THE GROWTH OF THE OPERATOR CONSTRUCT

	< 10	$10 \leq a < 11$	$11 \leq a < 12$	$12 \leq a < 13$	$a \geq 13$	Age Median	Age of Accession	Number of Girls	Number of Boys	Total Both
VII		GP 	BM BM		GP GP	12+	> 13	3	2	5
VI		BM	BM	GP BM	GM	12+	> 13	2	3	5
V		BM 	BP BP BP	BM BM 	GM BM GP BM GP BP	13	> 13	3	9	12
IV		GM BP BP BP BP	GM BM	GP BP GP BP GM BP BP	GM BM GM GP GM GP	12+	11+	10	10	20
III	BM BP	GM BP GP	GP GP	GP		10+	10+	5	3	8
II	GM BM GP BP	GM BM GM BP GM	GM GP	GM		10, 6	< 10	8	4	12
I	GM BP GP BP BM	GM BP GP BM	BM			10 10	< 10	4	6	10

Table 6 presents a record of the numbers of subjects classified at the various levels. Subjects are identified by representation faced (M or P) and by sex (G or B). The two experimenters studied each subject's record and classified subjects independently; there was 100% agreement on the classification of subjects. The table also gives the median age group for each level and the age of accession—the age group at which over 50% of the group are at that particular stage or above. For example, the age of accession for level IV is $11 \leqslant a < 12$, because for that group (and those older) 50% are at or above level IV, while less than 50% of the next younger group is at level IV.

Table 6 confirms the general pattern of performance gain with age discussed earlier. Yet at each age level, particularly the middle three, there is a wide distribution of performances. This is of particular interest because it is at these age levels that the great preponderance of fractional number instruction takes place. It is also interesting to note that in each performance category, the ages of the subjects vary considerably. There is

one girl aged 10 who mastered all tasks and one boy aged 11 who mastered none. Though this distribution occurs, it is useful to note that over 90% of subjects aged 12 and over fell in levels IV and above, while 74% of children below age 11 were classed below level IV.

The performance of the high achieving grade 10 students lends further validity and credence to the tasks and the category system. All of these subjects (their performances have not appeared in any previous analysis) were classified at level VII and their mean performance score was 66.5.

Further consideration of the nominal analysis above suggests that there are three general phases of performance or operator constructs. The first of these represents a rather primitive fractional construct and can be associated with "1/2 only" performance. Vestiges if this construct are seen in the "1/2 errors" which continue to be made on advanced tasks by subjects up to the age of 13 in this sample. The second operator construct is that of the unit operator and represents control over unit operators in general, including composition or multiplication of such operators. It was found that over 90% of the students who mastered "1/3 × 1/2" also could name it with the usual response being "divide by 6". The third phase of the operator construct is a general operator notion which incorporates unit and non-unit operators and compositions.

Thus Level II in Table 6 represents the "1/2 operator." Level III is transitional and leads to the "unit operator construct" represented by Level IV. Level V is a primitive general operator transition phase leading to the "general operator" construct. Level VII represents performance which is based on abstract manipulation of operators and may represent a fully general construct.

Sex Differences

It was not anticipated that there would be sex differences in performance on operator tasks. Thus, sex was not considered a sampling variable. Nevertheless the random sample contained 35 females and 37 males. The means for males was 34.3 and females 28.9; this difference was not significant (t = 1.05).

However, perusal of Table 6 suggests that 10 of 14 general operator phase performances by boys occurred under the machine representation while 6 of 8 general operator performances by girls occurred under the pattern representation. This sex by representation "interaction" is seen in Table 7. The sex difference under the machine representation is significant (t_{32} = 2.98) at the .01 level while the difference under the pattern mode is not (t_{36} = 1.11).

TABLE 7
MEAN SCORES BY SEX AND REPRESENTATION

	Machine	Pattern
Male	42.76 (n = 17)	27.15 (n = 20)
Female	24.41 (n = 17)	33.78 (n = 18)

Discussion

How or in what patterns do young persons think of fractions or rational numbers as operators? Both the analysis of mean performances and the nominal analysis suggest a pattern of growth over time. There appear to be three major phases in the operator construct, the "1/2 operator" phase, the unit operator phase, and the general operator phase. With transitions, the nominal analysis suggested 7 levels of performance across the ages 9-14.

What mechanisms does a young person use in coming to grips with rationals as operators? Noelting (1978b) has clearly shown several relevant phases in proportional thought. Particularly Noelting's phase IIa (p. 9) involving the equivalence class of the 1 to 1 ratio is sufficient to account for behaviour with "1/2 operators." His age of accession to this stage (8, 1) seems to relate well to age of accession to the 1/2 operator construct.

How does a young person gain a broader construct of fractional operators. While "equivalence class" thinking is germain to all 14 tasks in this study, it would appear that the mechanism of "partitioning" plays a central role as well. Partitioning is defined here to mean the division of a set into subsets of the same size and manifests itself in subjects' "division" statements. Thus a subject correctly pairs 2 with 6 in a "1/3 task" and explains "divides by 3." This kind of response was prevalent in this sample beyond age 10, with younger students more frequently giving subtractive explanations for their predictions.

The passage from the unit operator level to the general operator phase involves a more complex use of the mechanism of partitioning. For example a successful subject (age 13) pairs 90 with 60 in the 2/3 task and says "divide 90 by 3 and then multiply by 2" recognizing that whatever divides the input element 3 times divides the output element twice. An unsuccessful subject (age 10) realizes partitioning is involved but manipulates her explanation to fit her erroneous answers:

$$90 \rightarrow 50$$

"Three goes into 90 and 5 goes into 50."

It appears that control of the general fractional operators requires the partitioning of two sets or numbers but also coordinating this partitioning or, as Noelting (1978b, p. 15) suggests, using covariational thinking.

This covaried partitioning strategy occurred mainly in the subjects under the machine representation. In the pattern representation, a pattern explanation frequently accompanied a correct response. Thus in pairing 24 with 16 in the 2/3 task the subject would say "Well I knew 12 went to 8 so I just doubled it and got 16."

In some respects, seeing the pattern, particularly in non-unit fractional operators, represents more sophisticated thinking than the more rule-oriented covaried partitioning. This might explain the fact that the machine group reached a higher level of performance at a younger age than did the pattern group. Yet this use of partitioning also led to some clumsy if correct strategies as exemplified by this explanation "Well you divide 100 by 4 and then subtract that from 100." The high achieving grade 10 student who gave such an explanation also found great initial difficulty in handling non-unit inverse tasks. She never did reveal her final procedure but would say "I

would just think, what number do I divide by 4 and then subtract to get 60?" Although this strategy is plausible, her final response speed indicated some other procedure was in effect.

It would appear that students tested under the two representations used different strategies particularly with respect to non-unit operators. The machine representation appeared to induce a focus on a rule relating input to output, this rule most often involving covaried partitioning. The pattern representation was best handled by a recourse to known results; thus $24 \rightarrow 18$ is seen as $3(8 \rightarrow 6)$ and seemed to require knowledge, general proportional reasoning or, as Noelting sees it, "Common denominator thinking." It would be of significant theoretical interest to administer the operator tasks and the Noelting test to the same subjects and test the above conjecture. It would appear that young persons can cope with operator settings, at least in the machine mode, by "substituting" the mechanism of partitioning for a more sophisticated control over ratios.

An alternative explanation of the mode by age interaction concerns the notion of equivalence class. A subject who says "divide by 4 and subtract" sees the various "3/4" items as being determined by a single rule. In the case of a subject who sees $(8 \rightarrow 6)$ as $1/2(16 \rightarrow 12)$, the idea of equivalence as a class of ordered pairs which fit together appears to prevail. The former explanation of an item simply involves "working it out" while the latter appears to involve at least an implicit idea of a set of ordered pairs, the latter being a more formal mathematical notion as well as a more algebraic notion. It might be hypothesized that the machine mode sponsors the application of partitioning which in turn permits the use of a more primitive notion of equivalence in dealing with fractions as operators.

Nelson (in press) has suggested that "real" problems allow children to exhibit behaviours which seem beyond their predicted level of thought. Thus the "machine" representation might be an appropriate introduction to the operator construct with the less playful pattern representations and standard symbolic representations coming quite a bit later. To an extent this introduction to fractions obtains in various German curriculums today for students of about age 11. This age corresponds well with the results of Noelting (1978a) and this study. It would appear that a useful prerequisite for the study of general operators would be many different experiences with partitioning, particularly of discrete sets, as well as extensive work with unit fractional operators.

Before acting on the above curricular suggestions several questions need answering. From this study and to a much greater extent from the results of the previous study (Nelson & Kieren, 1978), it would appear that work with general rational operators would best be reserved for the junior high school. Is this the case, or can non-unit rationals as operators be successfully taught earlier?

It would appear that representation affects both attainment of the operator construct and the mechanisms used by young persons in building up the construct. To what extent is this true? What impact does development in the Noelting sense and instructional representation have on this result? Finally within this context, is the sex by representation interaction generally supportable? If so, what are the curricular impacts of this effect?

There appear to be three phases in the growth of the fractional operator construct. Moreover, there appear to be discernible developmental and constructive mechanisms used by young persons in building up this idea. The relationships of the operator subconstruct to other rational number subconstructs need to be illuminated and relevant learning experiences for young persons aged 8 to 15 devised.

References

Dienes, Z. P. *The elements of mathematics,* New York: Gordon and Bream, 1971.

Kieren, T. E. The rational number construct—Its elements and mechanisms. In T. E. Kieren (Ed.,), *Recent research on number learning.* Columbus: ERIC/SMEAC, in press.

Kieren, T. E., & Nelson, L. D. The operator construct of rational numbers in childhood and adolescence—An exploratory study. *The Alberta Journal of Educational Research,* 1978, *24*, 22-30.

Kieren, T. E., & Southwell, B. A comment on function composition. Edmonton: University of Alberta, 1978.

Lovell, K. Some aspects of the growth of the concept of a function. In M. Rosskopf, L. Steffe, & S. Taback (Eds.), *Projection cognitive development research and mathematical education.* Reston, VA: NCTM, 1971, 12-33.

Nelson, L. D. A hopeful trend in problem solving research. In. R. Lesh (Ed.), *Applied problem solving.* Columbus: ERIC/SMEAC, in press.

Noelting, G. *La construction de la notion de proportion chez l'enfant et l'adolescent et les mécanismes d'équilibration.* Québec: Ecole de Psychologie, Université Laval, 1978. (a)

Noelting, G. Constructivism as a model for cognitive development and (eventually) learning. Québec: Ecole de Psychologie, Université Laval, 1978. (b)

Padberg, F. *Didaktik der Bruchrechnung,* Freiburg: Herder, 1978.

Piaget, J., Inhelder, B., & Syzminska, A. *The child's conception of geometry.* New York: Basic Books, 1966.

Thomas, H. L. The concept of function. In M. Rosskopf (Ed.), *Children's mathematical concepts.* New York: Teacher's College Press, Columbia University, 1975, 145-172.

The Alberta Journal of Educational Research *Vol. XXV, No. 4, December, 1979*

B. BARBARA ADAMS
and
LLOYD O. OLLILA
The University of Victoria

The Relationships of Language Concepts, Auditory Comprehension, Visual Perception, and Spatial Relations as Predictors to Reading Achievement in First Grade

The investigators developed indices to measure some aspects of auditory comprehension, visual perception, and basic language concepts at the reading readiness level in order that new knowledge might be added to the construction of tests which are more accurate predictors of reading achievement in the first grade. Multiple linear regression analysis was performed to determine which subtests of the reading readiness tests were the best predictors of reading achievement. Results show that the Learning Rate subtest of the Canadian Readiness Test *and Spatial Relations Concepts of the* Experimental Reading Readiness Test *were highly significant predictors of reading achievement on all three subtests of the* Bond-Balow-Hoyt. *The sex differences in reading readiness were also investigated. The* Experimental Reading Readiness Test *indicated no significant differences between boys and girls in reading readiness levels. On the* Canadian Readiness Test, *however, the girls were superior on two of the subtests, Word Matching and Semantics. In reading achievement on the* Bond-Balow-Hoyt: The New Developmental Reading Tests, *there were no significant differences between first grade boys and girls.*

Some factors which educators feel contribute to reading readiness are auditory comprehension, auditory discrimination, richness of verbal concepts, general mental ability, knowledge of numerical and qualitative

Barbara Adams is an Educational Psychologist specializing in the field of Learning Disabilities and Reading Problems. She received her Master of Arts Education from the University of Victoria and is presently employed by the Greater Victoria School District.

Lloyd O. Ollila is Associate Professor of communications and social foundations in education at the University of Victoria, Canada. He received his B.A. and Ph.D. degrees at the University of Minnesota. His current research is in the area of children's word recognition, relationships between writing and reading, and reading readiness skills.

relationships, the sensory-motor abilities required in hand writing, the ability to attend, and the ability to follow directions (Johnson, 1969).

Standish (1959) and MacGinitie (1969) concluded from their review of the literature that the best predictors tend to be those tasks that are the most similar to the criterion, in other words a test of reading achievement. MacGinitie (1969) gave the following reasons for the predictive validity of tasks that resemble reading. First, if the predictor task is similar to reading, abilities required for the success in reading are measured by the predictor, even if it is not clear just what those abilities are. Second, predictor tasks measure both the environmental and motivational factors which are likely to encourage a child to become a better reader. Parental interest and encouragement will foster continued progress in the development of reading skills. Also the child who is curious enough to inquire about printed words will most likely continue to do so.

Many studies have shown that letter names are the best predictors of reading achievement. In order to name a letter correctly, a child must be able to discriminate visually one letter shape from another. Gates and Bond (1936), Barrett (1965), and de Hirsch, Jansky and Langford (1966) found that visual discrimination of words and letters was one of the best predictors of reading success.

The relative value of a child knowing the letter names, however, has been questioned by some researchers. Samuels (1971) and Ohnmacht (1969) argued against the importance of letter naming and claimed that letter sound was the more relevant variable. Ohnmacht demonstrated that children trained in letter sound performed significantly better than those trained in letter name. Many educators have jumped to the conclusion that because letter name knowledge is one of the best predictors of reading success that this should be taught before teaching the child to read. Barrett (1965) suggested that the ability to recognize the letters of the alphabet may be partially a reflection of past experiences with a variety of written materials which enable these children to recognize letters. These experiences may also help these children become better readers later. Barrett warned that one should not infer from his study that teaching children to recognize letters by name will necessarily ensure success in beginning reading.

These investigators are of the same opinion as MacGinitie (1969), Spache & Spache (1971), and Weintraub (1967) who concluded that test instruments must be found that are better predictors of reading achievement than those now in existence. The first grade teacher must know which children are ready for the instructional methods and/or the materials which she will be using in her classroom. For the children who are not ready, she must know their individual needs so that special readiness programs and materials can be prepared. Since it is not economically feasible to test all first graders individually, we must find group tests that are better predictors of reading readiness and are more useful as diagnostic tools for instruction. Research on reading readiness frequently fails to deal adequately with methods and materials for which children are required to be ready (MacGinitie, 1969). Goodacre (1971) stated that in the past readiness tests have been crude measures only, and therefore, identify only the extremes. In other words they could show only who was reading and who was not

reading. She also concludes that what happens to those children who have not started to read will depend on the instruction they receive from their teacher and her ability to diagnose the children's instructional needs.

Considerable re-examination of measures of reading as predictors of reading achievement is needed. Attempts should be made to define and measure children's developmental factors and to isolate and describe the functions in the reading process. Reading is a process in which the child's developmental characteristics assume varying importance at different stages.

Null Hypotheses

To investigate the remaining purposes of this study, the following null hypotheses were formulated:

1. On the investigator's *Experimental Reading Readiness* subtests and on the *Canadian Readiness* subtests, there will be no significant difference for sex.
2. On the *Bond-Balow-Hoyt: The New Developmental Reading Tests* (Bond, Balow & Hoyt, 1965), there will be no significant difference for sex.
3. On the *Experimental Reading Readiness Test,* the *Canadian Readiness Test* (Evanechko, Ollila, Downing & Braun, 1970), the *Bond-Balow-Hoyt: The New Developmental Reading Tests,* there will be no significant correlation between the various subtests.
4. On the *Experimental Reading Readiness* subtests and the *Canadian Readiness* subtests, there will be no significant predictors of reading achievement on the three subtests of the *Bond-Balow-Hoyt: The New Developmental Reading Tests.*

The Purpose of the Study

The primary purpose of this study was to develop new indices to measure the level of reading readiness of beginning first graders. It was not the authors' purpose to duplicate the reading readiness subtests that are already found to be good predictors of reading achievement. One area of auditory comprehension, visual perception, spatial relations, and language were considered. The specific aims were:

1. to identify and develop indices of reading readiness in the following areas: (a) Predicting Outcomes, (b) Basic Language Concepts, (c) Spatial Relations Concepts, and (d) Auditory Reception.
2. to determine which are the best reading readiness predictors of reading achievement,
3. to determine if there are significant differences in reading readiness and reading achievement scores between boys and girls on the tests used in this study.

Method

The Sample

Subjects used in this study were one hundred and fifteen children from seven first grade classrooms randomly selected from schools located in the Sooke School District on the lower Vancouver Island in British Columbia.

Sixty-two boys and fifty-three girls from rural and urban areas were involved in the investigation.

Contents of the Experimental Reading Readiness Test

The four following subtests were developed and administered to the first graders in September, 1975:

Test One: *Basic Language Concepts*—This 30-item test is designed to measure the child's basic language concepts or understanding of words which he has acquired through experience.

Test Two: *Predicting Outcomes*—This 30-item test is designed to measure the child's auditory comprehension and his ability to predict what will happen next when the context is given orally by the examiner.

Test Three: *Spatial Relations Concepts*—This 30-item test is designed to measure the child's concepts of space and his experiences and understanding of words, such as over, under, beside, and biggest.

Test Four: *Auditory Reception*—This 30-item test is designed to measure the child's ability to derive meaning from verbally presented statements and the ability to respond with a simple nonverbal *yes* or *no* by circling the word *yes* and the *happy face* if the statement is correct or the word *no* and the *sad face* if the statement is not correct.

Contents of the Canadian Readiness Test

Test One: *Technical Language of Literacy*—This 16-item test is designed to measure the child's knowledge of technical terms used in describing language, such as, letter, word, and numeral.

Test Two: *Letter Recognition*—This is a visual discrimination test of 21 items. It is designed to measure the child's ability to recognize both upper and lower case letters.

Test Three: *Word Matching*—This is a 21-item visual discrimination test.

Test Four: *Beginning Sounds*—This 19-item test is designed to measure the child's ability to discriminate between sounds at the beginning of words.

Test Five: *Semantics*—This 12-item test requires the child to classify stimuli by circling the three pictures, from a group of five, that belong together.

Test Six: *Learning Rate*—This subtest of 22 items is designed to measure the child's ability to learn sight words. After the teacher teaches ten sight words, the child is required to select from a row of three words one of the specific words named by the teacher.

Contents of the Bond-Balow-Hoyt:
The New Developmental Reading Tests

Part I: *Word Recognition*—This is a 36-item four-choice test designed to assess the pupil's word-recognition ability. The pupil is required to read the four words and mark the one that matches the picture.

Part II: *Comprehending Significant Ideas*—This 40-item test is designed to measure the child's ability to comprehend ideas expressed in a short paragraph.

Part III: *Comprehending Specific Instructions*—This 32-item test is designed to measure the child's ability to comprehend specific instructions which become more difficult as the child progresses through the test.

Administration of the Tests

The authors' *Experimental Reading Readiness Tests* (ERR) were administered by the classroom teachers during the second and third weeks of September, 1975. No more than two subtests were administered in any one sitting. The *Canadian Readiness Tests* (CRT), by Evanechko, Ollila, Downing and Braun (1970), were also administered during the same two weeks. During the last two weeks of February, the *Bond-Balow-Hoyt: The New Developmental Reading Tests* (BBH) were administered. All tests were scored by the teachers and then rechecked by the investigators.

TABLE 1
MEANS AND STANDARD DEVIATIONS ON THE *EXPERIMENTAL READING READINESS, CANADIAN READINESS,* AND *THE BOND-BALOW-HOYT: THE NEW DEVELOPMENTAL READING TESTS* OF FIRST GRADE CHILDREN

Variable	Boys (n=62)		Girls (n=53)		Both Sexes	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Experimental Reading Readiness Test						
1. Basic Language Concepts	24.98	2.44	25.58	2.40	25.26	2.42
2. Predicting Outcomes	22.68	3.66	22.26	3.46	22.49	3.54
3. Spatial Relations	22.98	3.53	23.42	3.60	23.18	3.54
4. Auditory Reception	22.45	5.18	22.23	4.02	22.35	4.64
Canadian Readiness:						
5. Technical Language of Literacy	13.21	3.37	13.15	4.12	13.18	3.70
6. Letter Recognition	17.87	3.95	18.21	4.10	18.03	3.99
7. Word Matching	15.26	4.63	17.09	3.75	16.10	4.31
8. Beginning Sounds	16.50	3.34	16.13	3.37	16.33	3.33
9. Semantics	27.76	7.50	30.45	6.77	29.00	7.23
10. Learning Rate	16.65	4.59	17.00	4.56	16.81	4.54
Bond-Balow-Hoyt: The New Developmental Reading Tests:						
11. Word Recognition	19.85	8.45	20.11	8.32	19.97	8.31
12. Comprehending Significant Ideas	15.53	10.84	17.40	10.30	16.39	10.55
13. Comprehending Specific Instructions	12.52	6.10	14.17	6.16	13.29	6.13

Results of the Study

The means and standard deviations are shown in Table 1 for the boys and girls on the *Experimental Reading Readiness Test, Canadian Readiness Test,* and the *Bond-Balow-Hoyt: The New Developmental Reading Tests.* On the *Experimental Reading Readiness Test* the means ranged from 25.26 out of 29 items (87.1%) on the Basic Language Concepts subtest to a low of 22.35 out of 30 items (75.0%) on the Auditory Reception subtest.

Sex Differences in Reading Readiness and Reading Achievement

The girls scored significantly higher than boys on only two subtests. They were the Word Matching subtest, ($T(113) = 2.29$, $p < .025$) and the Semantic subtest ($T(113) = 2.01$, $p < .05$) of the *Canadian Readiness Test.* All other differences between sexes on the subtests were not significant, $p > .15$ in each case.

Reliability of the Independent Variables as a Test on the Experimental Reading Readiness Tests

The Kuder-Richardson 20 reliability coefficient, which determines the internal consistency of the tests was computed for each of the four subtests on the Experimental battery. Coefficients of the four subtests ranged from .51 to .81. Two of the subtests, Auditory Reception and Spatial Relations Concepts, show a sufficiently high reliability of over .70 (Brunig & Kintz, 1968). The KR-20 reliability coefficient of .88 for the total *Experimental Reading Readiness Test* shows a high reliability when the *Experimental* battery is used as a whole.

TABLE 2
MEANS, STANDARD DEVIATIONS, RANGES, AND THE KUDER-RICHARDSON RELIABILITY COEFFICIENT FOR THE TOTAL SAMPLE ON THE *EXPERIMENTAL READING READINESS* SUBTESTS

Variable	Mean	SD	Range	KR-20
X1 Basic Language Concepts	24.69	2.51	17-29	0.505
X2 Predicting Outcomes	21.78	3.94	7-29	0.679
X3 Spatial Relations	23.03	3.58	13.29	0.706
X4 Auditory Reception	22.16	5.10	0.30	0.814
Total Test	91.65	11.40	50-112	0.876

N = 115

The Correlation Analysis

The Pearson Product-Moment Correlation Analysis was conducted to answer hypothesis 3, to determine the extent of the relationships among reading readiness and reading achievement measures. Table 3 shows the

TABLE 3

INTERCORRELATIONS BETWEEN SUBTESTS ON *EXPERIMENTAL READING READINESS, CANADIAN READINESS,* AND *BOND-BALOW-HOYT: THE NEW DEVELOPMENTAL READING TESTS*

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13
1.	1.000												
2.	0.372***	1.000											
3.	0.415***	0.466***	1.000										
4.	0.338***	0.349***	0.436***	1.000									
5.	0.281**	0.475***	0.358***	0.185*	1.000								
6.	0.331***	0.307***	0.341***	0.245**	0.536***	1.000							
7.	0.331***	0.432***	0.365***	0.259**	0.372***	0.453***	1.000						
8.	0.432***	0.418***	0.439***	0.362***	0.507***	0.621***	0.402***	1.000					
9.	0.449***	0.397***	0.398***	0.237*	0.415***	0.342***	0.407***	0.330***	1.000				
10.	0.357***	0.363***	0.362***	0.233*	0.462***	0.663***	0.590***	0.660***	0.310***	1.000			
11.	0.322***	0.338***	0.427***	0.264**	0.285**	0.405***	0.384***	0.454***	0.248**	0.574***	1.000		
12.	0.294***	0.220*	0.365***	0.240**	0.229**	0.316***	0.391***	0.358***	0.241**	0.465***	0.799***	1.000	
13.	0.348***	0.285**	0.426***	0.289**	0.262**	0.389***	0.434***	0.444***	0.303***	0.546***	0.834***	0.850***	1.000

* $p \leq .05$
** $p \leq .01$
*** $p \leq .001$

Readiness and Reading Tests (Readiness: Numbers 1 to 10; Reading: Numbers 11 to 13)

1. Basic Language Concepts
2. Predicting Outcomes
3. Spatial Relations Concepts
4. Auditory Reception
5. Technical Language of Literacy
6. Letter Recognition
7. Word Matching
8. Beginning Sounds
9. Semantics
10. Learning Rate
11. Word Recognition
12. Comprehending Significant Ideas
13. Comprehending Specific Instruction

intercorrelations for the entire sample. A two-tailed test was used to find the *p* significance. All the correlations proved to be significant; however, they varied as to the degree of correlation. *The Bond-Balow-Hoyt: The New Developmental Reading Tests* showed very high correlations between subtest (r = .80 to r = .85), suggesting that to a large degree they were measuring the same facet of reading achievement. The *Canadian Readiness* subtests showed lower correlations (r = .32 to r = .66). The correlations between the *Experimental Reading Readiness* subtests were lower (r = .35 to r = .47) showing a low degree of interrelatedness. One may assume that to a large degree each subtest is measuring a different facet of reading readiness on the *Experimental Reading Readiness Test.*

Multiple Regression Analysis to Find the Best Predictors of Reading Achievement

Hypothesis 4 was tested by developing and comparing multiple linear regression models to determine which readiness measures were the best predictors of reading achievement. The stepwise regression analysis was employed and the predictor variables entered the regression equation in order of their greatest contribution to the increase in R^2. Tested for significance was each entering variable.

Results showed that the Learning Rate subtest on the *Canadian*

TABLE 4
MULTIPLE REGRESSION—*EXPERIMENTAL READING READINESS, CANADIAN READINESS,* AND *BOND-BALOW-HOYT: THE NEW DEVELOPMENTAL READING TESTS* (SUBTEST 1)

		Subtest #1		
Step No.	Variable	Word Recognition % of Variance R^2	Probability Level	Standard Error
1.	Learning Rate	33.56	0.000*	6.839
2.	Spatial Relations	39.12	0.002*	6.576
3.	Predicting Outcomes	39.37	0.494	6.591
4.	Technical Language of Literacy	39.67	0.463	6.605
5.	Basic Language Concepts	39.85	0.570	6.625
6.	Auditory Reception	39.93	0.707	6.652
7.	Beginning Sounds	39.98	0.759	6.680
8.	Semantics	40.01	0.829	6.710
9.	Word Matching	40.01	0.973	6.742
10.	Letter Recognition	40.01	0.993	6.774

N = 115

TABLE 5
MULTIPLE REGRESSION—*EXPERIMENTAL READING READINESS, CANADIAN READINESS,* AND *BOND-BALOW-HOYT: THE NEW DEVELOPMENTAL READING TESTS* (SUBTEST 2)

		Subtest #2		
Step No.	Variable	Comprehending Significant Ideas % of Variance R^2	Probability Level	Standard Error
1.	Learning Rate	21.99	0.000*	0.396
2.	Spatial Relations	26.48	0.010*	9.162
3.	Word Matching	27.55	0.204	9.137
4.	Basic Language Concepts	27.91	0.458	9.155
5.	Predicting Outcomes	28.36	0.412	9.169
6.	Auditory Reception	28.59	0.554	9.196
7.	Letter Recognition	28.74	0.634	9.229
8.	Technical Language of Literacy	28.77	0.844	9.270
9.	Beginning Sounds	28.78	0.868	9.314
10.	Semantics	28.79	0.904	9.358

N = 115

Readiness Test was the best predictor of reading achievement on all three subtests of the *Bond-Balow-Hoyt: The New Developmental Reading Tests.* The Spatial Relations subtest on the *Experimental Reading Readiness Test* was also a significant predictor of reading achievement ($p < .01$ to $p < .001$). A second significant predictor of the *Bond-Balow-Hoyt*: subtest Comprehending Specific Ideas within the *Experimental Battery* was the Basic Language Concepts subtest. However, this occurred only in the absence of the *Canadian Readiness Test* from the predictor set ($p < .025$).

Discussion

The investigator's *Experimental Reading Readiness Test* had a high coefficient of reliability as a total test, indicating that it could be useful in testing for reading readiness, particularly if included with other testing.

Very few significant differences were found in reading readiness and reading achievement between boys and girls. There was no significant differences in reading readiness on the investigators' *Experimental Reading Readiness Test.* The girls, however, showed a significant difference in reading readiness on two of the subtests on the *Canadian Readiness Test,* Word Matching, and Semantics. There were no significant differences in reading achievement between boys and girls.

The subtest Spatial Relations Concepts on the *Experimental Reading*

TABLE 6
MULTIPLE REGRESSION—*EXPERIMENTAL READING READINESS, CANADIAN READINESS,* AND *BOND-BALOW-HOYT: THE NEW DEVELOPMENTAL READING TESTS* (SUBTEST 3)

		Subtest #3		
Step No.	Variable	Comprehending Specific Instructions % of Variance R^2	Probability Level	Standard Error
1.	Learning Rate	30.33	0.000*	5.162
2.	Spatial Relations	36.38	0.001*	4.954
3.	Word Matching	37.17	0.240	4.945
4.	Basic Language Concepts	37.80	0.295	4.943
5.	Technical Language of Literacy	38.21	0.398	4.949
6.	Auditory Reception	38.52	0.461	4.960
7.	Semantics	38.66	0.620	4.977
8.	Beginning Sounds	38.80	0.624	4.995
9.	Predicting Outcomes	38.91	0.668	5.014
10.	Letter Recognition	38.93	0.848	5.037

N = 115

Readiness Test proved to be the best predictor of reading achievement on all three subtests of the *Bond-Balow-Hoyt.* Basic Language Concepts was a significant predictor on the one subtest of the BBH-Comprehending Specific Instructions again, however, in the absence of the *Canadian Readiness Test.* It appears that both these subtests would be valuable contributors on measurements to predict reading. On the *Canadian Readiness Test,* Learning Rate proved to be a very strong predictor of reading achievement on all three subtests of the BBH. On this subtest the teacher teaches ten sight words. The child is then required to select from a row of three words one of the specific words named by the teacher. This subtest is designed to measure the child's ability to learn sight words, a very major part of learning to read.

From the findings of this study, Learning Rate and Spatial Relations Concepts appear to be important contributors to the prediction of reading readiness. Other variables, however, should not be discarded too soon. If the difficulty level of the *Experimental Reading Readiness Tests* had been increased and reliability improved, it is possible that the other subtests may have made a significant contribution as predictors of reading achievement.

Girls did not prove to be superior to boys in reading achievement, as had been found by Dykstra and Tinney (1969) where girls were found to be

superior in reading achievement. This may not have remained true if the reading tests had been administered in June instead of February. Further investigation into differences in reading readiness and reading achievement between boys and girls is required. Instructional needs of boys and girls should also be investigated.

More longitudinal research in reading achievement should be undertaken to clarify the abilities necessary for beginning reading and improve instruments to measure these abilities. Another area for research would be analyzing what readiness skills are important in a beginning reading program using an initial synthetic approach to word recognition as compared with a programme based on a whole word approach. Further research would also be useful in developing readiness tests which diagnose the best modality for the child's initial reading instruction.

References

Adelman, H. S., & Feshbach, S. Predicting reading failure: Beyond the readiness model. *Exceptional Children,* 1971, *37* (5), 349-354.

Barrett, T. C. Visual discrimination tasks as predictors of first-grade reading achievement. *The Reading Teacher,* 1965, *18*, 276-282.

Bond, G., Balow, C., & Hoyt, C. *Bond-Balow-Hoyt: The New Developmental Reading Tests.* Chicago: Lyons & Carnahan, Inc., 1965.

Brunig, J. L., & Kintz, B. L. *Computational handbook of statistics.* Glenville, Illinois: Scott Foresman, 1968.

de Hirsch, K. Tests designed to discover potential reading difficulties at the six-year-old level. *American Journal of Orthopsychiatry,* 1957, *27*, 566-576.

de Hirsch, K., Jansky, J., & Langford, W. *Predicting reading failure.* New York: Harper & Row, 1966.

Dykstra, R., & Tinney, R. Sex differences in reading readiness—First grade achievement and second grade achievement. In A. J. Figurel (Ed.), *Reading and realism.* Newark, Delaware: International Reading Association, 1969, 623-628.

Evanechko, P., Ollila, L., Downing, J., & Braun, C. *Canadian Readiness Test.* Unpublished, University of Victoria, 1970.

Gates, A. I., & Bond, G. Reading readiness: A study of factors determining success and failure in beginning reading. *Teachers College Record,* 1936, *37*, 679-685.

Goodacre, E. *Children learning to read.* London: Routledge & Kegan Paul, 1971.

Johnson, R. E. The validity of the Clymer-Barrett Pre-reading Battery. *The Reading Teacher,* 1969, *22*, 69-75.

MacGinitie, W. H. Evaluating readiness for learning to read: A critical review and evaluation of research. *Reading Research Quarterly,* 1969, *4*, 396-410.

Ohnmacht, D. D. The effects of letter-knowledge success—First grade. Paper presented at the annual A.E.R.A. Convention, Los Angeles, 1969.

Popham, W. J., & Sirotnik, K. A. *Educational statistics: Use and interpretation.* New York: Harper & Row, 1973.

Samuels, S. J. Letter-name versus letter-sound knowledge in learning to read. *The Reading Teacher,* 1971, *24*, 604-608; 662.

Spache, G. D., & Spache, E. B. *Reading in the elementary school* (2nd ed.). Boston: Allyn and Bacon, Inc., 1971.

Standish, E. J. Readiness to read. *Educational Research,* 1959, *12*, 29-38.

Weintraub, S. Readiness measures for predicting reading achievement. *The Reading Teacher,* 1967, *20*, 551-558.

The Alberta Journal of Educational Research *Vol. XXV, No. 4, December, 1979*

BETTY HORODEZKY
The University of British Columbia

Comparative Difficulty of Beginning Reading Vocabulary: Set II

A second set of 50 most frequently occurring words found in early primary reading materials was taught to 108 kindergarten pupils through the use of one of three instructional methods: the phonic, the kinesthetic, or the meaningful context. The rank order of difficulty for the 50 words was determined through ease of learning indices in the form of percentages based on the total number of correct responses recorded for each word. Data derived through use of the Spearman rank order correlation coefficient indicated that a nonsignificant relationship existed between rank order of word difficulty indices and rank order of frequency of occurrence for the words when taught by the three methods.

This study also included rank order indices for 13 additional words replicated from the first set of most frequently occurring words previously studied by Mangieri in an earlier investigation of this series. The rank order of difficulty indices reported in the present study for the same words was found to be not significantly different.

Numerous reading authorities agree that learning to read requires the child to develop an adequate sight vocabulary of meaningful words prior to instruction in word analysis (Chall, 1967). Despite those who express divergent opinions, the advantages of acquiring an early sight vocabulary are apparent for a number of reasons (Smith, 1963).

First, sight vocabulary development is necessary because it serves to facilitate the learning of difficult words which do not lend themselves to regular phonic analysis or structural word attack methods; second, sight words permit the child to focus on meaning rather than on "decoding"; third, there is greater immediate intrinsic value in learning to read whole words instantaneously than in learning to read letters in preparation for

---

Betty Horodezky received her Ph.D. in Reading and Language Arts from the University of Pittsburgh, where she served as Research and Teaching Associate. She has taught on the faculties of universities in Alberta, Colorado, Pennsylvania, and Nova Scotia. Dr. Horodezky has contributed to various IRA publications. She is currently a member of the Reading Department at the University of British Columbia.

reading (Anderson, 1968); and fourth, some sight words serve to provide basic understandings in phonetic and structural analysis which later enable the child to attach sound and meaning concepts to many unfamiliar words encountered in reading (Gray, 1960).

Regardless of whether a sight word approach or a word analysis approach is employed, initially, a decision concerning appropriate word selection for beginning reading must be made. Because of this need, numerous primary grade vocabulary lists have been developed over the past five decades. Although the authors of these works have rendered a significant service to reading, several major issues appear to have been overlooked in their quest to develop these lists. First, few significant studies were developed to investigate the relative learning difficulty of individual words contained within a given word list. Second, until recently only a small number of studies had been conducted to determine which sight words should be introduced first.

The present study, the second in a series of word studies, was thus initiated for the major purpose of more closely examining these problems (Sartain, et al., 1968, 1975). This investigation was designed to determine the relative leaning difficulty of a second set of 50 most frequently occurring words found in textbooks and library books for early primary level children, and to reassess the relative learning difficulty of 13 additional words previously studied by Mangieri (1972) and Mangieri and Sartain (1973), in a first set of 63 words taken from the Comprehensive Reading Vocabulary Primary Level (CRV-PL) list (Sartain, 1968-1975). The 13-word replication was done to ascertain consistency of findings. This investigation also included studies of the relationship between word frequency according to the CRV-PL list, the rank order of learning difficulty for the 50 words, and the relative effectiveness of the three methods of vocabulary instruction.

Method

Subjects

The population of the study consisted of 199 kindergarten children. Since the study was based on the acquisition of unfamiliar sight words, it was necessary that the subjects included in the investigation be limited to those who had not yet learned to read. The Slosson Intelligence Test (SIT) was administered to provide a stratified random sampling whereby subjects were placed into I.Q. categories of high, medium, or low ranges based on their respective SIT scores. Through random selection from within each I.Q. category, nine equivalent groups of 12 children each (4 high, 4 medium, and 4 low I.Q. range subjects) were provided to form a total sample of 108 subjects.

Through random selection, the 63 sight words were divided into three lists (A, B, and C) each containing 21 words. Three instructional methods—the phonic, kinesthetic, and meaningful context—were used to teach these words to the nine equivalent groups of subjects randomly assigned to Sections I, II, and III.

Procedure

Each section of three groups studied one word list of 21 words on a schedule of 7 words per week. In six weeks, each of the specified groups in a

given section was thus taught 21 words from either word list A, B, or C, utilizing one of three designated methods.

Each sample group of 12 was split into 2 groups of 6 in order to facilitate instruction. Teaching sessions consisted of 15-minute periods on 2 consecutive days in which seven words were taught and retaught by one of the three prescribed methods. Tests of recall were administered immediately following each instructional session and a delayed recall test was given approximately 24 hours after the second instructional session. All testing and teaching instruction was conducted by the writer who had extensive teaching experience at the primary level.

Analysis

The rank order of difficulty for the 63 words was determined through ease of learning indices in the form of percentages. These were based on the total number of correct responses recorded for the three tests of recall for each word. A comprehensive rank order index was conducted by totalling pupils' correct response scores for each word tested following instruction by all methods. The Spearman rank order correlation coefficient was used to determine the relationship between word difficulty and word frequency for the words studied. A comparison was made between the frequency of each word selected from the CRV-PL list, and the total difficulty index.

The effectiveness of the three instructional methods in the acquisition of word lists A, B, and C was assessed in a block treatment analysis of variance design. Post hoc comparisons were made through application of Scheffé's Method of Multiple Contrasts.

In order to determine if there was consistency in the rank order of difficulty indices findings between the first study and the present study for the 13 words replicated, the Spearman rank order correlation coefficient was applied. Table 1 provides data on the comparative rank order of learning difficulty indexes for these words.

TABLE 1
COMPARATIVE RANK ORDER OF LEARNING DIFFICULTY INDEXES FOR THIRTEEN WORDS REPLICATED FROM STUDY I

Present Study	Study I
Rank Order Ease of Learning for Thirteen Words Replicated	Rank Order Ease of Learning for Same Thirteen Words
1. as	1. see
2. look	2. look
3. see	3. are
4. down	4. down
5. at	5. all
6. and	6. she
7. she	7. as
8. not	8. not
9. are	9. at
10. all	10. and
11. how	11. how
12. what	12. then
13. then	13. what

Findings and Conclusions

It was found that a nonsignificant relationship existed between the rank order of word difficulty indices and rank order of frequency of occurrence for the 50 words when taught by using the three methods. These results suggest the frequency of occurrence of words found in primary level reading materials is not related to the learning difficulty of the words. The phonic, kinesthetic, and meaningful context instructional methods were not found to be significantly different in effecting acquisition of word lists A, B, and C. Also, the rank order of learning difficulty for the 50 words was not significantly different for any single method, as illustrated in Table 2. It can be inferred from these findings that the relative learning difficulty of sight words is not significantly affected by teaching method. A moderate to high correlation coefficient was found to exist between the rank order of difficulty indices for the 13 words replicated in the present study with their rank order of difficulty indices reported in the first investigation of this series.

TABLE 2
RANK ORDER CORRELATION
BETWEEN FREQUENCY OF OCCURRENCE
AND WORD DIFFICULTY INDEXES FOR WORD SET II (50 words)

Variables	2	3	4	5
1. Phonic Rank Order of Difficulty Index	.77*	.67*	.89*	-.15
2. Kinesthetic Rank Order of Difficulty Index		.76*	.94*	-.05
3. Meaningful Context Rank Order of Difficulty Index			.88*	.05
4. Comprehensive Rank Order of Difficulty Index				-.06
5. Frequency of Occurrence Rank Order Index				

* Significant at the .05 level.

Implications for Teaching

It might be inferred from the results of this study that words most commonly appearing in early primary reading materials are not of equal difficulty; some words appear to be more easily learned than others, regardless of their frequency of occurrence or of the method by which they are taught, as shown in Table 3.

In view of these findings, it is suggested that authors of primary reading materials and classroom teachers engaged in beginning reading instruction place greater emphasis on the relative learning difficulty of words and on the order in which these words are presented. Selection of appropriate words according to their ease of learning indices might well provide for the initial success so vital to the beginning reader, for "Among educators, parents, or psychologists, there would be few dissenters to the propositions that 'getting the right start' in learning to read is of greatest importance" (Heilman, 1972, p. 27).

TABLE 3
COMPREHENSIVE RANK ORDER OF DIFFICULTY FOR FIFTY WORDS (SET II) INCLUDING THE THREE METHODS TAUGHT

Comprehensive Rank Order of Words Learned		Ease of Learning Percentage	Rank Order of Frequency of Occurrence--CRV-PL List	Comprehensive Rank Order of Words Learned		Ease of Learning Percentage	Rank Order of Frequency of Occurrence--CRV-PL List
1.	too	70.30	10	26.	where	22.45	37
2.	Mr.	70.00	26	27.	time	19.00	42
3.	cat	54.46	43.5	28.	going	18.27	34
4.	if	53.06	25	29.	tens	17.82	15
5.	house	52.48	36	30.	want	15.84	12
6.	back	49.00	14	31.	long	15.00	50
7.	her	48.00	5	32.	can't	14.42	46
8.	know	47.57	8	33.5	something	14.29	23
9.	help	41.84	27	33.5	from	14.29	1
10.	man	39.42	17	35.5	were	14.00	2
11.	by	39.00	39	35.5	some	14.00	28
12.	red	38.84	41	37.	find	13.27	13
13.	or	37.62	38	38.	has	13.00	35
14.5	things	37.00	48	39.	day	12.00	29
14.5	don't	37.00	49	40.	got	11.88	47
16.	two	33.66	31	41.	words	11.00	16
17.	ran	33.01	18.5	42.	tell	10.89	24
18.	name	33.00	4	43.	into	9.71	22
19.5	make	31.68	6	44.	just	8.74	7
19.5	good	31.68	11	45.	right	7.92	32
21.	saw	31.00	33	46.	each	6.80	21
22.	play	29.70	43.5	47.5	them	4.95	9
23.	write	26.00	40	47.5	put	4.95	20
24.	came	24.49	30	49.	ones	3.85	45
25.	many	23.76	3	50.	could	0.00	18.5

References

Anderson, V. D. *Reading and young children.* New York: The Macmillan Company, 1968.

Chall, J. *Learning to read: The great debate.* New York: McGraw-Hill, 1967, pp. 13-14.

Gray, W. S. *On their own in reading.* Glenview, Illinois: Scott Foresman and Company, 1960.

Heilman, A. W. *Principles and practices of teaching reading.* Columbus, Ohio: Charles E. Merrill Publishing Company, 1972.

Mangieri, J. N. *Difficulty of beginning sight vocabulary taught by three methods.* Unpublished Ph.D. dissertation. School of Education, University of Pittsburgh, 1972.

Mangieri, J. N., & Sartain, H. W. Difficulty of beginning sight vocabulary taught by three methods. *Child Study Journal*, 1973, *3*(3), 107-114.

Sartain, H. W., and others. Unpublished series of vocabulary studies, University of Pittsburgh, 1968, 1972, 1973, 1975.

Sartain, H. W. Unpublished Comprehensive Reading Vocabulary Primary Level (CRV-PL) List. University of Pittsburgh, Falk Laboratory School, 1968-1975.

Smith, N. B. *Reading instruction for today's children.* Englewood Cliffs: Prentice-Hall, Inc., 1963.

JOHN McARTHUR
Monash University

Teacher Socialization: The First Five Years*

The study reported here relates to later stages of a longitudinal project which attempts to replicate in an Australian context the findings observed in North America and England regarding teacher socialization during the early career phase. The study involves a large sample monitored over a 5- to 6-year period and uses as its primary measure of socialization the Pupil Control Ideology Form, which is now well established in educational research. The basic finding is that, after an initial period of trauma, teacher socialization enters a plateau phase wherein those teachers who remain in their chosen occupation commence a process of value internalization.

How influential are the first five years of teaching? The study reported here attempts to provide some answers to this question by focussing upon one phase of a longitudinal study of beginning secondary teachers in Victoria, Australia. Earlier phases of teacher socialization have been described in terms of (a) "anticipatory socialization" during the teacher training phase, and (b) "reality shock", the phenomenon which occurs when idealism gives way to reality during the initial period of full-time experience. Phases (a) and (b) have been examined in a number of professions—medicine (Becker & Geer, 1958; Hughes, 1956), law (Lortie, 1959), and nursing (Corwin, 1961) are examples. In an earlier article describing phase (b), McArthur (1978) reported observations of "reality shock" with a group of beginning secondary teachers in terms of a movement in their self-reported "custodialism" towards pupils. A sample of 534 beginning teachers responded to statements on pupil control contained in the Pupil Control Ideology Form (PCI Form, Willower, Eidell & Hoy, 1967). By comparison with this group, who had commenced their initial teaching experience in 1973, a smaller group of 30, who had undertaken teacher training at the

John T. McArthur, B.Comm., B.Ed., Ph.D., is a Senior Lecturer in Educational Administration, Monash University, Australia. His teaching and research interests are in fields of teacher socialization, pupil control, organizational climate, computer techniques in educational research, multivariate statistical analysis. Dr. McArthur has been engaged in teaching, research, and publication at the university level since 1969 in school curriculum and administration.

* The study is supported by financial assistance from the Educational Research and Development Committee, Canberra, Australia.

TABLE 1
COMPARISON OF PCI SCORES BEFORE AND AFTER INITIAL TEACHING EXPERIENCE

Groups		Pretest PCI		Posttest PCI		
	N	Mean	Standard Error	Mean	Standard Error	*t*
Hoy (1968)						
Teaching	58 [a]	51.97	1.05	54.91	1.03	2.87*
Nonteaching	39	47.31	1.55	47.05	1.28	0.27
McArthur (1973)						
Teaching	534	44.51	0.41	50.88	0.44	17.63***
Nonteaching	30	45.57	1.49	45.20	1.69	0.40

[a] To enable more useful comparison, only the secondary teachers in Hoy's study (i.e., half his sample) are included in this table.

Probability levels for *t* values

*** Significant at .001
** Significant at .01
* Significant at .05

same time as the 534 but had not taught during 1973, displayed no change in their attitudes to pupil control. The findings of this section of the study were supportive of those reported earlier by Hoy (1969). Comparative results of Hoy (1969) and McArthur (1978) are included as Table 1.

Similar observations of the "reality shock" phenomenon in beginning teachers have been reported by Day (1959), using the Minnesota Teacher Attitudes Inventory, and by Morrison and McIntyre (1967) using the Manchester N, R and T scales devised by Oliver and Butcher (1962).

The study reported here is grounded in occupational socialization theory and the associated literature on reference groups (Hartley, 1960; Hyman & Singer, 1968). The suggestion that beginning teachers feel the need to comply with subcultural values of the teaching profession in order to achieve reference group membership is a common one. As Waller (1932) says:

> The significant people for a school teacher are other teachers . . . . A landmark in one's assimilation to the profession is that moment when one decides that only teachers are important. (p. 389)

It is also frequently argued that concern with discipline and control of pupils is a dominant subcultural value of teaching. (Waller, 1932; Silberman, 1970). According to Silberman, "The most important characteristic schools share in common is a preoccupation with order and control" (1970, p. 122).

The Notion of Pupil Control Ideology

The instrument used in this study as a proxy measure of teachers' self-reported attitudes to pupil control is the PCI Form. Details of the construction and validation of this instrument are contained in the original

monograph. The PCI Form consists of 20 items requiring Likert-type responses, ranging from Strongly Agree to Strongly Disagree. Respondents' scores are a summation of responses to the 20 items, ranging along a continuum from a possible low of 20 (not custodial) to a possible high of 100 (very custodial). Typical items are: "It is important for pupils to learn to obey rules rather than that they make their own decisions"; "Pupils often misbehave in order to make the teacher look bad."

PCI scores of teachers in a wide variety of school systems have been collected and analysed in relation to a range of biographical and other characteristics or attitude measures (e.g., sex, age, teaching experience, dogmatism, innovativeness). Willower (1977) has recently reviewed a large number of studies which have used the PCI Form. It would seem that there is sufficient evidence to justify acceptance of the notion of pupil control as measured by the PCI Form as an important aspect of the teacher subculture.

Prior Observations of the Sample of Teachers in this Study

In the earlier published report of phase (b) of the longitudinal study upon which the paper is based, some observations were made concerning the group of beginning teachers used to provide base-line data for the study. These observations are presented again here in summarized form for comparative purposes.

During 1972 a sample of 800 beginning teacher trainees, representing roughly 50% of the anticipated intake of new secondary teachers to the government school system[1] for the 1973 teaching year, were surveyed. The survey was conducted late in the year, after all practice teaching had been completed. Of the original group of 800, 564 (70.5%) completed a follow-up questionnaire in 1973, about 8 months after the completion of their teacher training courses. The PCI Form was used as the major part of the survey instrument on both occasions. Of the 564 subjects who completed the post-training questionnaire in 1973, 534 had been teaching during the intervening period and 30 had not taught at the secondary school level for various reasons (further study, pregnancy, illness, tertiary teaching appointment).

As reported above, a major finding of the earlier publication on the first stage of this study was the difference between those who had taught and those who had not. This finding supported Hoy's work and suggested a possible relationship between treatment (having been exposed to a teaching situation between pretest and posttest observations) and effect (increase in their custodialism as measured by PCI scores from pretest to posttest).

Differences between subgroups of the samples at both pretest and post-test phases were examined in cross-sectional analyses: e.g., based upon sex (males tended to be more custodial than females on both occasions); teaching subjects (teachers of physical sciences and practical subjects, such as physical education and home economics, were much more custodial than teachers of humanities subjects who, in turn, were more custodial than those prepared for counselling and guidance in schools); and commitment to teaching as a career (committed teachers tended to be more custodial than those uncommitted to a career in teaching). There were also marked differences between the group means of trainee teachers based upon the eight secondary teacher training courses extant in 1972. In most cases, at

the posttest stage these between-group differences persisted in the same rank order, precluding any real conclusions or speculations regarding differential patterns of socialization based upon observable characteristics of a "fixed" type; i.e., sex, teaching subjects, training institution. The exception was commitment, which, unlike the "fixed" characteristics, had a built-in flexibility, or potential for change. Very pronounced changes in PCI scores (between 24 and 33% compared with an overall gain of 13%) were recorded in the case of two of the nine possible situations when the three categories ("strongly committed," "weakly committed," "uncommitted") at both stages, were cross-tabulated. Paradoxically, not only those who moved from a position of weak commitment before teaching to a strong position after initial teaching experience, but also those whose strong pretest commitment had been changed to weak commitment after teaching, recorded strong PCI gains between the two observations. Dr. Itai Zak, Tel Aviv University, who is conducting a large-scale longitudinal study of beginning teachers in Israel (Zak, 1979), involving use of the PCI Form, has suggested a psychological explanation for this apparent paradox. In a recent personal communication he proposed that less stable teachers (i.e. those of high psychological uncertainty about the best course of action to choose to satisfy psychological needs) tend to gravitate towards the custodial end of the PCI Form continuum.

The Extended Longitudinal Phase of the Study

The present longitudinal study covers a longer period than most relating to the early years of teaching. Hoy (1969) published details of a follow-up of his sample of beginning teachers in their second year of teaching. It is not known whether he has kept track of this sample as they entered the mid-career phase. Wiseman and Start (1965) reported a follow-up study of teachers in England after their first five years of teaching. Walker (1967) conducted a five-year study of teacher commitment using a group of Australian beginning teachers. Charters (1970) touched upon aspects of this career phase in his study of teacher survival. Mason (1961) conducted a major study of beginning teachers in the U.S.A. but his research was mainly confined to the first year of teaching experience. Ortiz (1978) reviewed the topic of mid-career socialization of educational administrators but provided few details of the "make or break" period of the first five years. Becker's (1957) study of the career patterns of the Chicago public school teacher is one of the few detailed analytical studies of this vital career phase.

The present study appears to be one of the few longitudinal projects being carried out that involves a sizeable cohort of beginning teachers who have been monitored from their preservice years through to a stage of relative occupational maturity/security within their chosen profession in a large educational system. The data that are reported below were obtained in 1978 from 456 of the original sample of 800 trainee teachers surveyed in 1972. These 456 respondents can be categorised as follows:

	(a) Teaching in 1978	(b) Not teaching in 1978
1. Those who completed the survey on all three occasions, 1972, 1973 and 1978	278	80
2. Those who completed the survey in 1972, could not be contacted or failed to respond to the 1973 survey, but completed the 1978 survey (this includes some teachers who resigned during the period but have since rejoined the government teaching service)	76	22
Subtotal of 1978 survey respondents	354	102
3. Those who completed the 1972 and 1973 surveys but could not be contacted or failed to respond to the 1978 survey	194	12
4. Those who completed the 1972 survey but have not participated in either the 1973 or 1978 surveys (This includes 1972 participants known to have died or taken up permanent residence outside Australia. Many of these 1972 participants have never been employed in the Victorian government school system according to official records)	20	118
Subtotal of nonrespondents to 1978 survey	214	130
Total sample of respondents to 1972 survey	568	232

For the purposes of this report, categories 1(a) and 2(a) are the most useful in terms of an "experimental" group. The matching categories of 1(b) and 2(b) (comprising those respondents who were not teaching in 1978) provide a basis for comparison, possibly a "control" group.

Certain features of the "experimental" and "control" groups seemed worthy of cross-sectional analysis. These are discussed briefly here prior to consideration of the major findings of the study.

One interesting, somewhat surprising finding of the study concerns the career patterns of the teaching respondents during the period 1973 to 1978. Respondents were asked for details of place and nature of their employment for each year from 1973 to 1978 inclusive. Of those 354 who were still teaching in 1978, a remarkably high proportion (31.9%, i.e., 113 teachers) had maintained unbroken tenure at their school of initial appointment in 1973. 26 teachers resigned during the period 1973 to 1978 but had rejoined the government teaching service and were once again employed in full-time teaching positions in 1978.

The sex ratio of teachers in the sample of 354 in the experimental group reflected a predictable shift over the 5-year period. In the original sample of 800 teacher trainees, there were 480 females (60%) and 320 males (40%). In the experimental group of 354 there were 171 females (48.3%) and 183 males (51.7%). The career orientation of female beginning teachers has been commented upon by both Mason (1961) and Walker (1967), that is, a significant number indicated that they regarded teaching as a contingent career, one which they could move into and out of during their lifetime.

Analysis of Results

Data were analysed using the techniques of one-way analysis of variance, for cross-sectional comparisons, and *t* tests for correlated samples, for longitudinal comparisons of group means.

Cross-sectional analysis

From a simple comparison of the PCI scores shown in Table 2, one would conclude that there was a significant difference between the group means of those who were teaching and those not teaching in 1978. The results support earlier findings contrasting the PCI scores of subjects at the "coal-face" with those not teaching or in administrative situations (Willower, Eidell & Hoy, 1967; Hoy, 1968, 1969; McArthur, 1975, 1978).

TABLE 2
PCI SCORES OF RESPONDENTS TO 1978 SURVEY

Groups	*N*	Mean	Standard Error	*F*
Teaching in 1978	354	52.76	0.53	14.45***
Not teaching in 1978	102	48.53	0.86	
TOTAL	456	51.86		

Probability levels for *F* values

*** significant at .001
** significant at .01
* significant at .05

Cross-sectional comparisons were made using three characteristics as bases for grouping: sex, teaching subjects, and expressed preservice commitment to teaching. These are shown in Table 3 along with comparative figures from the 1972 sample, both the total group of 800 and the specific group of 354 respondents who were teaching in 1978.

Earlier observations that males tend to display more custodialism in their PCI scores than females (Willower, Eidell & Hoy, 1967; McArthur, 1978) were confirmed by the findings shown in Table 3. The increased proportion of males in the 1978 survey was commented upon above. The generally accepted phenomenon of secondary teaching as a young, increasingly "female" profession, at least in the Australian context, is not denied. It is suggested however that global figures often conceal some interesting underlying patterns.[2]

Differences in mean PCI scores when groups were based upon teaching subjects were similar to those observed during the "reality shock" phase (McArthur, 1978). Teachers of sciences and practical subjects were much more custodial than those teaching humanities. Those few trained in guidance and counselling but actually teaching in classrooms in 1978 (n = 4) were still very "humanistic" compared with all other teachers. Most of the colleagues of teachers in this small group were now either in consultative positions with the counselling and guidance service centres of the educational authority or had left the teaching service for other employment.

TABLE 3
PCI SCORES OF TEACHING RESPONDENTS TO 1978 SURVEY, CLASSIFIED ON BASIS OF SEX, TEACHING SUBJECTS, AND PRESERVICE COMMITMENT TO TEACHING

Groups	*N*	Mean	Standard Error	*F*
Male	183	53.74	0.75	3.57**
Female	171	51.71	0.76	
Humanities subjects	181	50.97	0.75	
Sciences and Practical subjects	169	54.94	0.74	9.42***
Guidance subjects	4	42.00	4.02	
Strongly committed	186	52.79	0.74	
Weakly committed	94	52.47	1.11	0.08
Uncommitted	74	53.12	1.10	
1978 teaching respondents	354	52.76		

Probability levels for *F* values

*** Significant at .001
** Significant at .01
* Significant at .05

Preservice commitment was not found to be a useful predictor of differences in custodialism in the 1978 sample. Strength of commitment tended to be associated with slightly higher custodialism than weak commitment. Lack of commitment was associated with slightly higher custodialism than either of the other two categories. The differences between group means were far too small to make any further comments worthwhile.

Longitudinal Analysis

Data analysis in this section was confined to those 358 respondents to the 1978 survey for whom PCI scores had been obtained on each of three occasions (1972, 1973, 1978). While 278 of these were teaching in 1978, 80 had ceased teaching for various reasons: leave for travel overseas; further study; homemaking; pregnancy; secondment as specialist consultants; other occupations.

Table 4 shows comparative scores for the three surveys; the responses of those teaching in 1978 are separated from those not teaching at that time.

The plateau "effect" observed supports the basic notions of theory in occupational socialization. The "reality shock" phase is accompanied by dramatic shifts as idealism is replaced by realism. Thereafter, a flattening of the curve occurs as stability and security in the occupation consolidate attitudes.

Hoy (1969) observed a difference between those teachers who moved to a new school in their second year of teaching. In contrast to those who remained at the same school and whose PCI scores in Year 2 were

TABLE 4
COMPARISON OF PCI SCORES OF 1978 RESPONDENTS WHO ALSO RESPONDED TO 1972 AND 1973 SURVEYS

Groups	*N*	1972 PCI		1973 PCI		1978 PCI	
		Mean	Standard Error	Mean	Standard Error	Mean	Standard Error
Teaching 1978	278	45.45	0.58	52.24	0.61	52.72	0.61
Nonteaching 1978[a]	80	43.75	0.87	47.89	1.10	48.18	0.98
		F value 2.04*		*F* value 11.48***		*F* value 13.02***	
	358	45.07		51.27		51.70	

[a] A number of those not teaching in 1978 were teaching in 1973. This accounts for the apparent contradiction between these figures and Table 1. In fact only 18 of the 30 "nonteaching" respondents represented in Table 1 are included in the 80 "nonteaching 1978" in Table 5.

Probability levels for *F* values

*** Significant at .001
** Significant at .01
* Significant at .05

TABLE 5
COMPARISON OF PCI SCORES IN 1972, 1973, AND 1978 SURVEYS OF 1978 TEACHING RESPONDENTS, CLASSIFIED ON THE BASIS OF STABILITY OF TENURE

Groups	*N*	1972 PCI		1973 PCI		1978 PCI		Correlated *t* tests		
		Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	1972:73	1972:78	1973:78
Still at same school as in 1973	91	45.45	1.03	52.71	0.94	52.59	0.98	7.58***	6.31***	0.09
Taught in 2 or more schools	143	46.04	0.80	52.54	0.87	53.31	0.88	9.67***	8.07***	1.06
Resigned between 1973-77 but rejoined the service	44	43.52	1.50	50.30	1.75	51.05	1.64	3.98***	4.01***	0.75
		F value: 1.04		*F* value: 0.80		*F* value: 0.26				
Total	278	45.45		52.24		52.72		13.49***	11.57***	0.82

Probability levels for *F* and *t* values: *** significant at .001
** significant at .01
* significant at .05

practically identical with those in Year 1, the mobile teachers seemed to become more custodial. Hoy postulated a "resocialization" phenomenon as the underlying cause. As noted above a sizeable proportion of the experimental group in the 1978 survey were at the same school from 1973 to 1978 thus it was considered appropriate to subdivide the group on the criterion of stability of tenure. The 278 teachers in the experimental group

fell into 3 distinct categories for the purpose of this section of the analysis: those who had taught at 2 or more schools during the period; and those who had resigned but later rejoined the teaching service during the period. Comparative figures are shown in Table 5.

The trends shown here are interesting although not statistically significant. A surprisingly high proportion of teachers have remained at their schools of initial teaching appointment (91, or 32.7% of the 278 still teaching in 1978 and who responded to all three surveys) and the mean PCI score of this group is lower than that of the 143 who have taught at two or more schools in the same period. This high rate of stability of tenure suggests an interesting side issue with respect to career patterns. There may be some relationship between increasing teacher mobility and lack of promotion opportunities in the situation of teacher over-supply common in most Western industrialized countries. The planned follow-up interviews with approximately one-third of the 1978 survey respondents may throw further light upon this issue.

A further set of longitudinal analyses was undertaken using the three variables isolated for cross-sectional purposes in Table 3, namely, sex, teacher training subjects and expressed preservice commitment to a career in teaching. Subgroup PCI means for 1972, 1973 and 1978 are shown in Table 6. As reported in earlier discussions comparing the 1972 and 1973 data (McArthur 1975, 1978), the response patterns over time of subgroups classified on these three bases are not sufficiently dissimilar to allow claims for possible underlying causal factors.

TABLE 6

COMPARISON OF PCI SCORES IN 1972, 1973 AND 1978 SURVEYS OF 1978 TEACHING RESPONDENTS, CLASSIFIED ON BASIS OF SEX, TEACHER TRAINING SUBJECTS, AND EXPRESSED PRESERVICE COMMITMENT TO TEACHING AS A CAREER

Groups	*N*	1972 PCI		1973 PCI		1978 PCI		Correlated *t* tests		
		Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	1972:73	1972:78	1973:78
Male	143	46.41	0.88	53.07	0.85	53.38	0.83	8.70***	7.40***	0.49
Female	135	44.42	0.75	51.36	0.89	52.02	0.90	10.68***	9.17***	0.85
		F value: 2.92*		*F* value: 1.24		*F* value: 1.24				
Humanities subjects	141	43.48	0.81	51.21	0.84	50.78	0.86	10.82***	8.50***	0.54
Sciences and practical subjects	133	47.87	0.81	53.58	0.89	55.09	0.84	7.96***	7.61***	1.72
Guidance subjects	4	34.25	3.23	44.00	3.72	42.00	4.02	3.51*	4.31*	0.85
		F value=10.31***		*F* value=3.23**		*F* value=8.30***				
Strongly committed	147	45.98	0.81	52.54	0.82	52.25	0.85	9.55***	8.06***	0.40
Weakly committed	75	45.43	1.02	52.83	1.16	53.39	1.21	9.04***	6.50***	0.48
Uncommitted	56	44.07	1.42	50.56	1.47	53.04	1.32	4.69***	5.16***	1.60
		F value: 0.88		*F* value: 0.58		*F* value: 0.25				
Total	278	45.45		52.24		52.72		13.49***	11.57***	0.82

Probability levels for *F* and *t* values: *** significant at .001
** significant at .01
* significant at .05

Further analysis is being undertaken to establish whether differential patterns may be more discernible using multivariate rather than univariate procedures. Some interesting differences persist over time in the data when the 1972 training course is used as the basis for grouping respondents. Furthermore, there are some marked differences observable when these institution-based groups are further subdivided on the basis of sex.

Another area of continuing interest is the relationship between changes in expressed commitment to teaching as a career over a period of the three surveys and the extent and direction of changes in custodialism as measured by self-reported PCI scores in 1972, 1973 and 1978. This issue will be examined in a further report on the data, incorporating follow-up interviews.

Conclusion

The analyses reported in this article indicate that the most marked adjustment in the socialization of beginning teachers takes place during the period of initial teaching experience. This period of "reality shock" is followed by a consolidation or plateau phase during which those who choose to remain in the teaching profession adjust and "internalize" values of the teaching subculture. There is slight, perhaps insufficient, evidence in the data to suggest that the plateau is not so level for those who experience some instability during this phase of "midcareer socialization." Such instability may result from movement to a new school involving "resocialization," new identification, perhaps even compliance with a new reference group. Or it may result from moving out of teaching for a year or two then back into the profession.

As yet the data analysed are insufficient to enable firmer conclusions regarding these matters. Obviously an underlying bias is present in terms of self-selection of respondents who have chosen to remain in teaching. The high proportion of respondents shown in Table 3 as having expressed strong commitment (196, or 52.5% of the 354 teaching respondents in 1978) when compared to the corresponding number in 1972 (336, or 42% of the total of 800) bears testimony to this underlying bias. Despite this potential for distortion the results that have been analysed appear to support the notion that the post "reality shock" years are much less traumatic than the period of initial teaching experience.

Notes

1. The government school system accounts for about 80% of the pupils, teachers and schools at the secondary school level (years 7 to 12). In 1973 there were 243,800 pupils, 17,00 teachers, and 354 schools at the secondary level in the Victorian government school system.
2. For example, the median age of secondary teachers in Victoria is 28.9 years. Of the 2,930 secondary teachers under the age of 25, 2,089 (71.3%) are female. The age group 25 to 29 years comprises 5,541 secondary teachers, 3,060 (55.2%) of whom are female. Of the remaining 5,984 secondary teachers in 1978 who were aged 30 years and above, only 2,629 (37.9%) were female. The median age of 28.9 corresponds roughly to the "mid-career" phase, 5 to 6 years after training, that is the basic focus of this paper.

References

Becker, H. S. The career of the Chicago public schoolteacher. *American Journal of Sociology,* 1957, *63,* 470-477.

Becker, H. S., & Geer, B. The fate of idealism in medical school. American *Sociological Review,* 1958, *23,* 50-56.

Charters, W. W. Some factors affecting teacher survival in school districts. *American Educational Journal,* 1970, *7,* 1-27.

Corwin, R. The professional employee: A study of conflict in nursing roles. *American Journal of Sociology,* 1961, *66,* 604-615.

Day, H. P. Attitude change of beginning teachers after initial teaching experience. *Journal of Teacher Education,* 1959, *18,* 326-328.

Hartley, R. G. Relationship between perceived values and acceptance of a new reference group. *Journal of Social Psychology,* 1960, *51,* 181-190.

Hoy, W. K. The influence of experience on the beginning teacher. *School Review,* 1968, *76,* 312-323.

Hoy, W. K. Pupil control ideology and teacher socialization: A further examination of the influence of experience on the beginning teacher. *School Review,* 1969, *77,* 257-265.

Hughes, E. C. The making of a physician: A general statement of ideas and problems. *Human Organisations,* 1956, Winter, 21-25.

Hyman, H. H., & Singer, E. (Eds.) *Readings in reference group theory and research.* New York: The Free Press, 1968.

Lortie, D. C. Laymen to lawmen: Law schools, careers and professional socialization. *Harvard Educational Review,* 1959, *20,* 352-369.

Lortie, D. C. *Schoolteacher.* Chicago: University of Chicago Press, 1975.

Mason, W. S. *The beginning teacher: Status and career orientations.* Final Report. Washington, D.C.: U.S. Department of Health, Education and Welfare, 1961.

McArthur, J. T. *Teacher socialization: The influence of initial teaching experience on the pupil control ideology of beginning secondary teachers.* Unpublished doctoral dissertation, Monash University, 1975.

McArthur, J. T. What does teaching do to teachers? *Educational Administration Quarterly,* 1978, *14,* 89-103.

Morrison, A., & McIntyre, D. Changes in opinions about education during the first year of teaching. *British Journal of Social and Clinical Psychology,* 1967, *6,* 161-163.

Ortiz, F. I. Midcareer socialization of educational administrators. *Review of Educational Research,* 1978, *48,* 121-132.

Oliver, R. A. C., & Butcher, H. J. Teacher attitudes to education: The structure of educational attitudes. *British Journal of Social and Clinical Psychology,* 1962, *1,* 56-59.

Silberman, C. E. *Crisis in the classroom.* New York: Random House, 1970.

Walker, W. Occupational commitment in a sample of Australian teachers five years after training. *Australian Journal of Higher Education,* 1967, *3,* 20-32.

Waller, W. *The sociology of teaching.* New York: Wiley, 1932.

Willower, D. J. Schools and pupil control. In D. A. Erickson (Ed.), *Educational organization and administration.* Berkeley: McCutchan, 1977.

Willower, D. J., Eidell, T. L., & Hoy, W. K. *The school and pupil control ideology.* Penn State Studies Monograph No. 24. University Park: The Pennsylvania State University, 1967.

Wiseman, S., & Start, K. B. A follow-up of teachers five years after completing their training. *British Journal of Educational Psychology,* 1965, *35,* 342-361.

Zak, I. (1979) Developmental transitions in the professional growth of teachers. In *Teacher Effectiveness: Proceedings of the 4th Conference of the International Association of Educational Assessment.* New York: Pergamon Press, 1979.

The Alberta Journal of Educational Research *Vol. XXV, No. 4, December, 1979*

J. M. KIRMAN
and
J. GOLDBERG
The University of Alberta

Student Teacher Telephone Conferencing With Satellite Maps as a Monitoring Device*

This exploratory study was conducted to determine the feasibility of telephone conferencing for student teaching supervision. It was found that student teachers performed as well with this mode of supervision as those supervised face to face. However, the student teachers expressed dissatisfaction with the procedure.

As faculty budgets are cut back, the employment of additional staff for supervising student teachers becomes more difficult. The extended practicum of some jurisdictions also places additional demands on the time of faculty members. The use of telephone conferencing is one way of allowing faculty members to supervise larger numbers of student teachers from a central location, rather than travelling to several locations for supervising a limited number of students.

Present technology allows any number of individual telephones to be joined into a conference. Thus, a faculty advisor may link his student teachers in a conference call at a mutually agreed to time. Technological improvements are presently being developed that may enable a central university switchboard to set up conference calls at a cost comparable to ordinary telephone calls (Edmonton Telephones, Note 1).

In this project three criteria were employed in assessing the effectiveness of telephone supervision: teachers' ratings of student teachers; student

Dr. J. M. Kirman is a Professor in the Department of Elementary Education, Social Studies area, and director of Project OMEGA for research in remote sensing education.

Dr. Goldberg is an Associate Professor in the Department of Educational Psychology, and research associate with Project OMEGA.

* This project was funded by Innovative Projects Fund, Alberta Department of Advanced Education and Manpower, Program Services Division. The cooperation of Dr. Al MacKay, Director of the Center for Research in Teaching, University of Alberta, for external evaluation is gratefully acknowledged.

teachers' ratings of their supervision; and judgment based on examination of the knowledge of select pupils of the student teachers. The latter was based upon specially prepared social studies teaching units on Landsat satellite infrared false color maps.

The following questions were raised:

1. How would student teachers supervised with telephone conferencing compare to another group of student teachers supervised in a face to face manner as evaluated by their cooperating teachers?
2. How would pupils taught about Landsat maps by the above two groups compare with each other?
3. How would the student teachers in each group rate their experiences with the two modes of supervision?
4. What are some variables that influence telephone conferencing as a supervisory tool in student teaching?

Previous Research

There appears to be no literature specifically dealing with telephone conferencing in which a group of student teachers and their faculty advisor are on line at the same time. Some reports of telephone conferencing deal with a telephone receiver hooked to a loudspeaker (Stevenson & Otto, 1975). Another deals with conferencing lecture and discussion groups, but not with student teacher supervision (Madden, 1967). In one case, telephone conferencing was used but only with a single student, the faculty consultant, and the cooperating teacher at one time (Dalrymple, 1971). One study also made use of this "triad" procedure on a single telephone line. Amplifying devices were used by the parties to allow two or more persons to hear or speak with the same receiver (Smith, 1969).

The University of Wisconsin, Télé-Université, Québec, and the Open University, London, England, among others, utilize telephone communications for instruction. Usually, groups of students assemble at central locations for telephone loudspeaker lectures with two-way communication. London also has one to one telephone tutorials and group conferences (Williams, Holloway & Hammond, 1975).

The San Diego Schools utilize group telephone assisted instruction for handicapped children, requiring special installations at the students' homes (Carr, 1964). However, the group telephone project being reported here used existing telephone lines and equipment in the students' homes without any modifications.

Procedure

The study was run during both halves of the school year 1977-1978 with The University of Alberta, Faculty of Education Practicum 201. This is an experience in which participants are expected to teach small groups of pupils. It usually consists of five consecutive weeks where the student teacher is in a school for two days per week. Faculty advisors' roles consists mainly of orienting students and cooperating teachers, following the students' progress, and dealing with problems.

The use of Landsat satellite maps was decided upon to measure the

achievement of the student teachers' pupils. These maps are in infrared false color with a scale of 1:1,000,000. They are produced by Landsat satellites in a near polar orbit, 900 km high, which produces images of almost all the earth's surface every eighteen days.

Since this technology is relatively new, it was expected that neither cooperating teachers, student teachers, nor their pupils had any background regarding satellite maps. This provided a topic in which all student teachers began on an equal footing, and made the student teachers more dependent upon their faculty advisor.

Each student teacher had to teach three children, selected by the cooperating teacher, about Landsat infrared maps for one hour each week. The three children selected were those with the highest reading scores in the class. The children were in grades four, five, and six, and were checked for visual problems. Research had shown that children as young as those in grade three could deal with infrared Landsat maps (Kirman, 1977).

The student teachers were to meet before their teaching round for an orientation session. Telephone and control groups were scheduled to meet at different times, and were assigned to different schools. The subjects were 19 randomly selected student teachers, 10 of whom were randomly selected and allocated to telephone-conference supervision, and the remaining 9 allocated to face-to-face supervision. Table 1 provides a description of the subjects on four variables. The data are not available for one student teacher from the telephone group. However, all other data for this subject are available and are therefore included in all subsequent analyses.

The sex ratio of 2 males and 7 females was identical for both groups. Also, the figures for mean number of years at university (3.6, controls; 3.7, subjects) and mean number of years in the Faculty of Education (2 years in control; 2.3 years, subjects) were virtually identical for both groups. The mean age of the telephone group (25.1) was 3.2 years more than the mean age of the control group. A t test on the differences between the age means was clearly nonsignificant ($t = 1.17$; d.f. = 15) so the two groups can be regarded as comparable in age.

For the controls there were five weekly supervision sessions of at least 45 minutes in each of two schools over the five weeks of the project. The frequency and duration of these supervision sessions exceeded what is normally available to Education Practicum 201 students.

In addition to instruction about Landsat maps, the sessions for the controls dealt with topics pertinent to beginning student teaching. Topics

TABLE 1
COMPARATIVE DATA FOR STUDENT TEACHER GROUPS[a]
($N = 18$)

Group	Mean Years University	Mean Years In Education	Mean Age	Number of Male	Number of Female
Control	3.6	2	21.9	2	7
Telephone	2.7	2.3	25.1	2	7

[a] Data not available on one telephone subject.

were introduced both by the faculty consultant and by the students and handled in group discussion, unilateral presentation, or questions and answers.

The same faculty consultant supervised the telephone subject group in one conference hook-up of 45 to 60 minutes on Thursday evenings for the five consecutive weeks of the project. All participants were at home.

After greeting students and outlining the plan for the evening, the faculty consultant invited each to give a brief account of experiences in school that week. The faculty consultant responded to each account and/or invited others to respond. If necessary, the faculty consultant designated a particular person by name to respond.

Some time was always allocated to Landsat teaching. Also, students selected the previous week reported on topics such as discipline in their class or children with learning problems. The faculty consultant tried to promote group discussion on the presented topic. Time was also devoted to problems. Students with problems which could not be brought up in conference were invited to telephone the faculty consultant privately immediately after. The conference terminated with the faculty consultant providing goals for the student teaching and following conference session.

Note that the control group was in two schools, thus giving each other greater peer support, whereas the telephone group was scattered throughout many schools. The telephone group was usually limited to a maximum of forty-five minutes per week with their faculty consultant. A smaller ratio of student teachers to supervisor favoured the control group since separate 45-minute supervisory sessions were held for the controls in each of the two schools. In addition, the control group was not subject to the stringent time limitations applied to the telephone group. Since many features of the supervision favoured the control group, the present study would appear to be a severe test of telephone supervision.

Evaluation By The Cooperating Teacher—
The Cooperating Teacher Evaluation Form

The student teachers were evaluated by the cooperating teachers who employed the Cooperating Teacher Evaluation Form. This is a standard form requiring qualitative evaluative comments under the headings of attendance, punctuality, attitude, preparation, relationship with pupils, and general comments.

At most, student teachers could be separated into two groups on the basis of the cooperating teachers' ratings—those with consistently positive evaluative comments, and those with generally positive evaluations qualified by a reservation expressed about some teaching characteristics.

The frequency of consistently positive ratings was higher in the telephone group. Nine of ten telephone subjects received a consistently positive rating whereas only three of nine controls received a consistently positive rating. A chi-square test showed this difference to be significant at less than the .05 level ($\chi^2 = 4.33$; d.f. = 1). One can conclude that the telephone group received better global evaluation from their cooperating teachers than did the control group.

A possibility exists that the telephone group's mode of supervision may

TABLE 2
RATINGS ON COOPERATION TEACHER EVALUATION FORM
FOR TELEPHONE AND CONTROL SUBJECTS
(N= 19)

Rating	Group	
	Telephone	Control
Positive	9	3
Qualified Positive	1	6

$\chi^2 = 4.33$; d.f. = 1; $p < .05$.

have made them more dependent upon the cooperating teachers. This in turn may have fostered a closer relationship which could have had a bearing on the higher evaluations received by the subjects.

The question can also be raised as to whether either mode of supervision resulted in greater student teacher effectiveness than would have been produced with no supervision at all. This cannot be answered by the present data because an unsupervised control group was not included for two reasons: apprehension of this arrangement; and the newness of Landsat material requiring somebody to monitor the student teachers.

Assuming that supervision did make a difference, it can be concluded, at the very least, that telephone supervision would not appear to be poorer than traditional supervision in terms of student teacher effectiveness as assessed by cooperating teachers in this project.

Evaluation Of Pupil Achievement in Landsat Map Reading--
The Pupil Map Questionnaire (P.M.Q.)

The P.M.Q. consisted of 15 questions based on two Landsat maps of Alberta which the pupils perused in responding to the questions. One part, eliciting a total of 16 responses, required the pupil tested to point to various features, e.g., find a lake. As the child pointed, the examiner noted whether the response was right or wrong. The other part of the P.M.Q. focused on more general knowledge and permitted less objective scoring. In view of the differences between the two parts, they were separated for analysis.

The P.M.Q. was administered individually to the 30 pupils taught by the telephone supervised group and to the 23 pupils taught by the control group of student teachers.

As shown in Table 3 the pupils of the telephone group obtained a higher mean score on the 16-item section (12.9 vs. 11.87) requiring them to point out specific locations on Landsat maps. The difference between the means is statistically significant at the .05 level on a one-tail t test. On the appropriate two-tail test (since direction of difference was not predicted), t approaches significance ($p < .10$ level).

As seen in Table 4, all pupils achieved at virtually identical levels on the general questions (pupils of telephone group, 6.14; pupils of controls, 6.21). High means suggest that both groups may have attained their ceilings on

TABLE 3
COMPARISON OF PUPILS' ACHIEVEMENT ON
16-ITEM PUPIL MAP QUESTIONNAIRE
(N=53 pupils)

	Pupils of Telephone Ss	Pupils of Control Ss
$\bar{X}$	12.9	11.87
S.D.	1.87	2.11

$t = 1.68$; d.f. = 51
$p < .10$ (two-tailed test)

TABLE 4
COMPARISON OF PUPILS' ACHIEVEMENT ON
7-ITEM PUPIL MAP QUESTIONNAIRE
(N=51 pupils) [a]

	Pupils of Telephone Ss	Pupils of Control Ss
$\bar{X}$	6.14	6.21
S.D.	.901	1.099

[a] The responses of two of 53 pupils to the verbal questions were lost due to problems in tape recordings.

$t = .244$; d.f. = 49
N.S. (of two-tailed test)

this kind of test in Landsat knowledge. The t test on the difference between these means was nonsignificant ($t = .244$).

The data support the conclusion that telephone supervision is clearly not associated with a lesser degree of teaching effectiveness in student teachers.

Since supervision was the main source of Landsat teaching knowledge and techniques for the student teachers, the data offer some support for the conclusion that telephone supervision is at least as adequate as face-to-face supervision in transmitting knowledge about Landsat teaching.

Self-Evaluation of Supervision—
The Supervision Evaluation Form (S.E.F.)

The student teachers' evaluation of their mode of supervision was determined by their responses to the Supervision Evaluation Form (S.E.F.) which was completed anonymously after the final supervisory experience. The items covered aspects of the supervision ranging from its usefulness in providing feedback on their performance to its usefulness in stimulating more thought about student teaching. Content validation is the only method of validity determination for the S.E.F. This is assumed by the form's

construction which consists of systematically delineating functions supervision is intended to serve, and formulating items corresponding to each function.

All eight controls who returned their S.E.F. forms expressed acceptance (ratings of excellent, very good, or good) of face-to-face supervision, whereas only one of nine responding telephone subjects expressed overall acceptance of telephone supervision.

Only two of nine telephone subjects reported telephone supervision to be acceptable, whereas only one of nine would select telephone supervision, if given a choice. Responses indicated that if a choice had to be made between telephone supervision once weekly versus face-to-face supervision once or twice every five weeks, seven of nine subjects would prefer face-to-face supervision.

The results suggest that the telephone group perceived a deficiency in interaction with others during supervision and that the supervision was not sufficiently individualized. Other reasons for dissatisfaction included one expressed frequently about the "impersonal" nature of contact over the phone and the lack of "direct" contact: "It's nice to sit next to someone and ask them how it's going", "would be nice to see people", etc. Complaints expressed no more than once by a single person included poor timing of the call, and the unavailability of the home phone which was needed for business.

Most controls indicated they *would* reject telephone supervision even when the alternative was infrequent face-to-face supervision. Five of the controls spontaneously referred to the disadvantages of telephone supervision although they were not required to contrast their method with telephone supervision! One control noted, "It would be an unpleasant ordeal to sit with an ear to the phone for 45 minutes." These negative references seem remarkable since the controls did not experience telephone supervision, nor have contact with the telephone subjects. This suggests that the controls may have had preconceived negative attitudes toward telephone supervision, and that people generally, including the telephone subjects, may have shared this negative attitude. One of the challenges in using this supervisory mode, therefore, would be to overcome this attitude.

The telephone and control groups did not differ on a majority of the S.E.F. specific items dealing with objectives of supervision such as learning about Landsat, learning from others, obtaining new ideas, being stimulated to think. One can, therefore, venture the conclusion that telephone supervision was functional for imparting important information to student teachers in general, though it succeeded less well in responding to individual needs. Also, a majority of the telephone group listed as an advantage the convenience of telephone supervision.

Conclusions

Student teachers supervised by telephone conferencing were rated as performing better by their cooperating teachers than those supervised in face-to-face meetings with their faculty consultant. The pupils taught by the telephone supervised group did as well as the control group's pupils in learning about Landsat maps. The telephone group, however, reported much

dissatisfaction with telephone supervision. Since similar dissatisfaction regarding telephone conferencing was found at Open University, London as well, this must be considered as being a problem which has to be dealt with in future research.

Since telephone supervision was found to be associated with better or equal performance on two independent indices of student teacher effectiveness and since this method is more time-efficient and can be more economical, it is concluded that additional research on it should be carried out.

While the relationship depicted by the data permit several alternative explanations and hence cannot be regarded as providing unequivocal support for the claim that telephone supervision is more effective, or even for the lesser claim that it is *sufficiently effective,* the latter two claims are consistent with the data.

The study provides an empirical basis for pursuing further research on this method of supervision. In addition to probing the causal relationship between supervision and student teacher effectiveness, research should probe ways of making telephone supervision more palatable to the participants.

Questions For Further Research

Can more positive feelings toward telephone conferencing be generated by reducing the number of participants on line to allow more individual participation, limiting the time that each participant might speak, and providing thorough training for conference conductors?

Can one-way video used in conjunction with two-way telephone communication overcome feeling of impersonalness, by allowing the students to see their faculty consultant during the conference?

Is there a relationship between student teacher performance and the type of supervision by the faculty consultant?

Does a close relationship between the student teacher and cooperating teacher affect student teacher performance?

Notes

1. Edmonton Telephones, Meeting for discussion and demonstration of conference telephone bridging device, May 19, 1978.

References

Carr, D. B. Teleteaching is inaugurated for students who are homebound. *California Education,* 1964, *1,* 21-22.

Dalrymple, J. I. Remote supervision at preservice and inservice levels in teacher education—Research reported in the vocational education special interest group at the American Educational Research Association meeting, 1971 AERA Convention. ERIC #ED050032 SP004900.

Daniel, J. S., & Turok, B. Teaching by telephone, a two-way communication mode in distance education. *Proceedings of the 10th I.C.C.E. World Conference,* Brighton, England, May 1975, 133-140.

Kirman, J. M. The use of infrared false color satellite images by grades 3, 4, and 5 pupils and teachers. *The Alberta Journal of Educational Research,* 1977, *23,* 52-64.

Madden, C. F. Amplified telephone as a teaching medium; description of an inservice science seminar. In H. Bosley & H. E. Wigren (Eds.), *Television and related media in teacher education, some exemplary practices*. Baltimore, Maryland: Multi-State Education Project, 1967. ERIC #018978. EM006153. Article originally published in 1965.

Smith, P. M. *Experimentation to determine the feasibility of remote supervision of student teachers*. Unpublished doctoral dissertation, University of Ohio, 1969.

Stevenson, M. C., & Otto, A. C. Telephone teaching and testing: University of Nevada, Reno. *Journal of Home Economics,* 1975, *67,* 27-28.

Williams, E., Holloway, S., & Hammond, S. Students' reactions to tutoring by telephone in Britain's Open University. *Educational Technology,* 1975, *5*, 42-46.

PERSPECTIVES

Some Observations Regarding Value Added By Education

Rupert N. Evans
University of Illinois at Urbana-Champaign

Although economists and tax experts have long been interested in assessing the value added by manufacture, little attention has been paid to the value added by education. The value added to an individual by an educational program can be expressed in terms of the individual's annual income, score on a job satisfaction scale, level of socioeconomic status, or any one or combination of measures intended to assess the effects of education. The value (however it be determined) of persons who are entering an educational program can be subtracted from their value upon completion of the program, and the difference is the value added by that particular educational program. Or, the value of persons who have not had a particular educational program can be subtracted from the value of a similar group of individuals who have completed a part or all of the program. Again, the difference is the amount of value added.

At first glance, value added by education may seem to be only another way of stating a concept which is usually referred to as "gains." Educational researchers have long used pre- and post-test designs which measure gains made by one or more groups which have received different treatments. Teachers have compared gains in achievement when they used different treatments. In almost all such cases, however, the consideration of "gains" is incidental. There has been an assumption that groups were nearly equal at the start of experimentation, and "gains" has been used almost solely for increasing the precision of measurement by making minor corrections in the starting point of similar groups. The ultimate criterion used in judging the education of an individual, however, is almost invariably the value of that individual as a finished product. This note, on the other hand, maintains that it is more important to consider the gains

Rupert N. Evans has served on the faculty of the University of Illinois (Urbana, Illinois, USA) since 1950, serving as Professor, Associate Dean, and Dean of the College of Education. Currently he is Professor of vocational and technical education in the Bureau of Educational Research. He holds the Ph.D. (in industrial psychology) and an honorary doctorate in vocational education from Purdue University. He has written many books and articles on work and education.

which are made by each individual. The sum of these gains is "value added by education."

At the present time, every phase of education except special education is concerned almost exclusively with value of the student as a finished product. I have noted that professors who work with doctoral candidates try to select those students who will reflect glory on their advisor and his/her institution. Vocational educators seek students who will be most successful in their careers. Music teachers seek students who are already talented musicians.

Tests used in education are usually concerned with value of the student as a final product, rather than with value added to that student by education. I have observed students who have learned to fail advanced placement tests so that they can take elementary courses and thus obtain a few high grades. This enables them to survive in a school which is less concerned with value added by education than with value of the final product. Teachers in schools which require standard examinations have learned to teach for the tests and to fight for above average classes in order to survive in a system that measures value of the finished product instead of value added by education. Teachers in every school I have observed prefer elective, rather than required courses, in part for the same reason. Athletic coaches search for talented athletes, rather than for students whose athletic abilities can be developed markedly. Their coaching abilities are measured by final results, rather than by value added. The dramatic coach, the band director, and the advisor of the school newspaper are under similar pressures.

The employment managers of manufacturing companies can and probably should have as their goal the selection of the best possible employees they can hire at the rates they can pay. The goal of most companies is increased value added by manufacture, and they choose people who can help to achieve this goal. On the other hand, it seems to me that the goal of educational institutions should be increased value added by education, so they should choose teachers and plan programs which can aid in achieving this end.

Unfortunately, at present, the ideal students are those who already know everything they will be taught. They are the students who will be the best finished products, and reflect the greatest credit on their teachers and their school. The ideal student is thus the one who does not need the education he/she will receive; the student for whom education will add no value. Something must be wrong with this logic.

Vocational education offers some of the most interesting examples of the need to measure the worth of education by the value it adds. Since the passage of the Smith-Hughes Act in the United States in 1917, vocational educators in that country have followed the precept that vocational education should be offered only to those people who can succeed in the occupation for which they are being prepared. There is nothing wrong with this precept. Obviously, if a person is trained to become a commercial artist, but he cannot succeed because no one discovered that he was color-blind, very little value has been added by the vocational education he has received. It *is* important that he be able to succeed in the occupation for which he has been prepared.

Unfortunately, many vocational educators all over the world have warped this precept. If they have a choice of students, they select those who are *most likely* to succeed in the course for which they are being selected or in the occupation for which they are being prepared, rather than choosing students who *can* succeed. For example, students who are amateur radio operators and who have designed and built their own electronics equipment are often preferred students in an elementary level vocational electronics course, even though they almost certainly will learn less in that course than someone who knows no electronics. The ideal apprentice often is one who previously has acquired all of the skills of a journeyman. In too many cases, the ideal student is one who does not need the educational program. He or she is one who probably can succeed without additional training.

It seems obvious that in many cases the person who will profit most from education is the person who has ability which is relatively undeveloped. If she has a handicap which can be removed easily by education, that education will have added great value to her and to society. If a very brief educational experience serves to identify and make visible a person who would not otherwise be considered employable, that education has added a great deal of value. If an educational experience develops ability for self-education, it is likely to add considerably more value than an experience which ends with the end of formal instruction. If an educational experience makes a difference between employability and non-employability, its value added is likely to be greater than the value added by an education which provides somewhat higher income to a person who will be employable regardless of added education.

In certain cases, however, the best candidates for training really are the best candidates who apply. A talented musician may become even more skillful under the guidance of a skilled teacher. A skilled automobile mechanic may be the best candidate for a course in diagnosis of automotive malfunctions. But in each case, the test of the selection of students and the worth of the educational program should be on the basis of the value which can be added by the training program, not on the basis of "value of the product."

We rarely are in a position to measure directly the values added by education. There have been attempts to measure the value added to earnings by government scholarships, but they have been contaminated by a suspicion that those who chose to use a scholarship were probably more likely to have succeeded with or without it than those who did not choose to use this educational benefit. Similar doubts cloud the many attempts to measure the value added by a high school or a college education.

There is, however, at least one clear case where data are available on value added by education. In 1962, Dr. Solomon Chaneles began a program of vocational training at Ruykers Island in New York City. Ruykers Island is a correctional institution for youthful offenders. When these young men were committed, they were given a test used by IBM for selection of computer programmers. A random half of those who passed the test were trained as programmers under a grant from the U.S. Department of Labor. The other half went through the standard education program of the institution. Both groups had the same raw material, but the value of the

finished product was clearly superior, according to every measurement, for the group which received vocational training. Even in this case, no precise measurement of value added by education has been attempted.

The precise measurement of value added by education is a problem which needs to be tackled by the educational economist, the educational psychologist, the educational philosopher, and the educational sociologist. All four groups need to be involved, because we should be concerned not only with the value added in terms of dollars, but also in terms of personal worth and worth to society.

Precise measurement of value added by education would be desirable, but is unlikely to come soon. In the meantime, however, much can be accomplished if the administrator, the teacher, and the public can begin to think in terms of value added by education. If they could simply ask themselves, "Will this action of mine add to the value of education for John Jones or Sally Smith", we would move forward much more rapidly. Entrance requirements to various educational institutions would undoubtedly be changed if we thought in these terms instead of considering only the value of the finished product. Nursery education might pay more attention to children from lower class homes. Scholarship requirements might be changed. Less attention might be given to tax rebates and more attention paid to scholarships based on potential for growth.

Probably the greatest difference would occur in the classroom. There is possibly considerably greater marginal utility in more attention by teachers to the poorest and to the best students in a class than there is in a concentration on the average student. If a student can be taught enough to keep him off relief for most of the rest of his life, the value added by education is probably far greater than a similar gain in knowledge by the average student. Similarly, a small amount of extra effort with brilliant students may lead them to self-instruction which will add a great deal to their value and to the value added by their education.

ESSAY REVIEWS

Essay Review I

EDUCATION IN ATLANTIC CANADA

D. A. MacIver
University of New Brunswick

EDUCATION AND DEVELOPMENT IN ATLANTIC CANADA. *Edited by Eric Ricker.* Halifax, N.S.: Department of Education, Dalhousie University, 1978, 381 pp.

There are no urban areas in Atlantic Canada that are able to absorb large numbers of rural immigrants. At least there are no centers to compare with Vancouver, Calgary, Toronto, Montreal, or even Saskatoon. These large cities are expanding urban areas which have been able to absorb much of the migrating rural population. As centers of power (and as the "consumers" of rural and other immigrant populations), cities have dominated the educational scene in North America regardless of the myth of the one-room school. The reasons for this development will not be argued here. It is the results of this process from the perspective of education and economic development in the non-urban environment that will be considered. The opportunity to do this is provided by the record of a conference, *Education and Development in Atlantic Canada,* which was edited by Eric Ricker of Dalhousie University.

This collection of papers brings together, probably for the first time, a variety of issues which reveal the problems of education and rural development. Unfortunately, it takes some effort to extract the significant material from Ricker's document and this is because of the way that Ricker has defined his task. When an editor undertakes to publish the record of a conference, he has to decide whether to be selective or comprehensive. Ricker chose the latter course and his collection of papers and discussions represents the good and the bad, the banal and the superb. This choice undermines the usefulness of the volume because some excellent contributions are hidden amongst a few tedious articles and some rambling and misdirected discussion. As a conference report this collection does its job. It is less successful as an analysis of the relationship between education and

---

Don MacIver has been a teacher, principal, university professor and Dean. He has edited two books on Canadian education and is currently concerned with the problems of non-urban education in Canada. Dr. MacIver has worked in Western, Central, and Eastern Canada, and has been Dean of the Faculty of Education, University of New Brunswick for the last seven years.

development in Atlantic Canada because the lack of selection disturbs the focus of the volume.

Having established these reservations, anyone interested in education and development in Atlantic Canada will be rewarded by slogging through the collection. Ricker's introductory article sets the stage and introduces a theme that is picked up in the better articles. The theme is that the schools in the Canadian East were never designed for the non-urban environment. From the 1920's on, there is evidence of a desire to centralize schools in order to provide an urban system of education. The blighted economic condition which exists in Atlantic Canada is due in no small part to the imposition of this urban concept of education. Unfortunately, it seems that neither the schools nor the universities have been able to provide an alternative that meets the needs of non-urban communities. Young people have drifted from the rural areas to the local urban communities which have never developed sufficiently to absorb all the rural population. The migration continues from the local centers to the cities of central Canada which have been forced to accept the people of the Atlantic region as immigrants.

The Atlantic Canadian has had a particularly bleak existence and the future is uncertain. Local entrepreneurs and multinationals have failed to establish a typical urban economy east of Montreal. So the Atlantic Canadian has been caught between a declining rural economy and an urban economy that has never recovered from a delicate infancy. These failures, so goes the argument, are compounded by the colonial mentality that pervades the East. This mentality has so undermined the confidence of the people that they are unable to see the potential that currently presents itself. Compared with urban communities, Atlantic Canada does not have to deal with major problems of pollution, crime, overcrowding and renewal projects. There seems to be a better future available, but habits of the past inhibit the imaginations.

Instead of seeking to capitalize on the indigenous resources, the politician and businessman in Atlantic Canada persist in trying to establish urban industries, much as an addled boxer reacts to a bell long after his fighting days are over. This is not unique, for almost nowhere in the world has education yet responded to the concept of non-urbanization (or re-ruralization) that seems to be in embryo in Atlantic Canada and elsewhere. More than one author wonders if the institution of education can respond to these new and sophisticated life styles. If teachers and professors remain as unresponsive as they have been in the past, then education is likely to become even less relevant to Atlantic Canada than it is at the moment. This is the general theme introduced by Ricker and pursued by the critical papers in this collection.

John Saul's papers considers the concept of Atlantic Canada as a colony of central Canada and other industrialized regions of North America. Schools "reinforce bourgeois domination and depersonalize the working classes," making them docile tools whose ambition is to be like the bourgeois (p. 74). There is the need to develop indigenous analyses and indigenous solutions to local problems, but the colonial tendency always to look elsewhere is encouraged by the educational system. Saul cites Tanzania as a country that is trying to protect itself from the "global ripoff" by

developing its own self-reliant economy. But this approach, Saul argues, is reflected in the system of education and there is a confidence in the people that seems to be missing in Atlantic Canada.

Saul's neo-Marxist critique seems to be accepted by David Macdonald. Macdonald is a harsh critic of the schools and universities. He sees them "as mindless purveyors of inappropriate goods, (and) our educational system has failed Atlantic Canada," but he goes on the claim that however witless, the schools remain the main hope. Again, like Saul, Macdonald cites the socialist Nyerere as a model leader, a man who is using education to build the confidence of his people, a confidence that the elite of Atlantic Canada do not have in their people. Macdonald accepts the apocalyptic theories of Illich and argues for a much less comprehensive educational system. While urging a diminution of the control of formal education, he argues for a much greater involvement of the community in education. One cannot help but think of Dewey's theories as one continues to read Macdonald. The paper gradually softens from the radical views of Illich to reflecting the later writing of Dewey in which the schools become continually more associated with the community. Clearly this is not a new idea, but perhaps its time has come.

Another of the more refreshing papers in the collection is written by Linden McIntyre, a veteran Halifax journalist. He argues that the system of education has been "criminally incomplete because it has not been responsive to local and regional situations" (p. 158). Again, one reads that the Maritimes have been treated as a colony and that industrial development has failed. What is needed is "small primary-based units . . . even if subsidized and protected by government" (p. 153). The theme that some third world countries have found some answers and that socialist policies may be useful is implied in this article as in the works of Macdonald and Saul.

If socialist ideas can improve the economic conditions of Atlantic Canada, it is not unreasonable to wonder about the absence of any significant radical party in Atlantic Canada. The interlude in this volume that deals with the Antigonish Movement suggests some answers to this question. Father Mifflin provides a careful description of this religious, yet broadly based social movement, and his comments are elaborated in the succeeding articles and discussions. The days of the Movement were exciting and stimulating even if "times were hard." The excitement did not always arise from overcoming the social conditions and fighting the bosses. Such activity occurred, but there seems to have been even more activity of a kind that was fatal to the development of a real socialist movement.

Sacouman argues that trade unionism (and by implication, socialism) did not develop in Nova Scotia due to the lack of "proletarianization." The typical "bluenoser" was a jack of all trades, a part-time miner, a part-time farmer, a part-time lumberman, and so on. When the pits closed he went back to the farm and did not join a labour union, but he was attracted to the Antigonish Movement which was a social force led by dynamic priests.

By the very nature of the leadership, the Movement played it cosy with authority. Coady, a founding father of Antigonish, was a friend of cabinet ministers and had a "harmonious relationship with the government" (p.

184). Indeed, Lotz points out that the government of the day assisted the Movement and as a result

> It took a lot of weight off the back of the political system. If there had not been an Antigonish Movement, we might today be about twenty years ahead in terms of political reforms. (pp. 175-76)

The Antigonish Movement, it seems, undermined the development of socialism and other reform movements in Nova Scotia.

There were also problems within the Movement itself. The activist priests met with resistance from within the Church while the bishop of Antigonish, a supporter of the Movement, antagonized the bishop of Halifax, who was not. Then it seems that there was a special position adopted by the Scottish Catholics who became embroiled with the university . . . and so it went on. That, however, was only *within* the Church.

Social reform was undermined by religious differences. An occasion is cited when the United Church invited some representatives of the Antigonish Movement to a meeting. After a brief discussion all the Lutherans and Anglicans walked out leaving the Movement representative speaking to an almost empty hall. And then, of course, there were socialists and trade unionists whose efforts seem to have been emasculated by the Antigonish Movement. The Antigonish Movement overwhelmed the trade unions and socialist parties of the 30's and 40's and gradually the Movement diverted radical reformist tendencies into educational enterprises.

With a suddenness that jars, this fascinating discussion of the Antigonish Movement is dropped and the reader is confronted by a radical treatment of a contemporary issue. Edgar Friedenberg writes on the role of the schools in Atlantic Canada and his paper is crammed with insights. This paper is clearly and succinctly written, though one does not get that impression from the discussions that follow. Fortunately, it is not necessary to read the record of the discussion to get Friedenberg's point. In essence, he claims that the schools of Nova Scotia are the worst schools in Canada:

> In the degree to which they discourage and make problematical for kids even the degree of self direction within the approved limits of choice which is essential for the most modern kind of acculturation, with all of its disadvantages. (p. 201)

Friedenberg repeats the assertion that Atlantic Canada is a colony and that schools train children to be colonized or to leave and head for central Canada. He emphasizes that there are various ways of perceiving the Maritimes and he enjoys the Nova Scotian lifestyle while recognizing its limitations. He has no desire to see Atlantic Canada become a "developed country," that is:

> One that is in such command of its technical resources that it has the economic power to shift its industrial waste—human and otherwise—to elsewhere. (p. 213).

But there is an alternative that attracts Friedenberg and that is simply the idea of capitalizing on the local environment.

Friedenberg, for all his sophistication, is not so far from the views of many of the invited participants. They believe that education ought to have a good deal to do with improving the lives of people in the places where they live with their families and friends. Education ought to be concerned with many options and not merely that of becoming "developed," whether one

uses Friedenberg's idiosyncratic definition or not. Indeed, "development" is precisely *not* the question for Friedenberg. A common criticism becomes apparent. It is that few Atlantic Canadians have benefited from an education that has enabled them to participate fully in an urbanized and bureaucratized society. But most Atlantic Canadians are unable to see the possibilities of their environment because educators, functioning in response to the dominant pressures of society, are also unaware of the possibilities.

The theme of a misdirected approach to education is in Jim McNiven's paper. McNiven is an economist closely associated with industrial development in Atlantic Canada. This is not to argue that McNiven agrees with Friedenberg, because they take different positions and deal with different issues, but the fact remains that both believe that a more relevant approach to education is necessary in Atlantic Canada. McNiven argues that society has become overwhelmingly urbanized and bureaucratized which is not entirely a bad thing, even though society has been caught in the megamachine of bureaucratic control. He writes:

> For most people throughout most of history, the adoption of these skills and values would have been a small price to pay for the stability and material standards that the megamachines offer. The advantages cannot be questioned. (p. 275)

With all its advantages, however, there is a move to "re-ruralization," that is, a movement of people back to the country because of dissatisfaction with urban existence.

McNiven takes great care to establish that the rural environment is not what it was. It is an environment that has been urbanized even though the population is sparse. That is, communication and organization have made isolated communities dependent on the contemporary social network. The school has been responsible for much of this kind of modernization and at the same time the centralized school has been responsible for the depopulation of the countryside. The structure of the school imposes an acceptance of a bureaucratic lifestyle which is another characteristic of urbanization. It is not possible to return to an idyllic past, but the schools do not prepare the young for the possibilities of a non-urban future.

As well as the lifestyle that one would anticipate, McNiven indicates that contemporary society has become a "single-crop society." He claims that if one large sector of our contemporary urban society were to collapse, the whole complex might settle about our ears. This contrasts with a rural society where, even under the most adverse conditions, whole societies may suffer calamities from drought and disasters of like kind and yet survive. McNiven's implication is that if we continue pursuing education as it exists today, disaster is inevitable. Only by seeking alternatives, especially rural alternatives, will it be possible to have some confidence that disaster may be averted.

A paper by Smitheram and another by Tilley reveal once again the urban orientation of schools but the perspective is different and McNiven's paper is nicely complemented, especially by Smitheram's work. Tilley's paper is a historical review of the decline of community control of schools in Nova Scotia. One of the interesting insights revealed in the paper is that before 1940 there was probably a lot less autonomy than one may have imagined.

Both of these papers contribute to a general picture of the decline of the rural school in Nova Scotia and Prince Edward Island.

The implication of this collection of papers is that the theory and practice of education in Eastern Canada has been urban-oriented for many years. This has probably been the case for education in the rest of Canada, but there is an important difference. That is, from Victoria to Montreal the process has worked *with* the prevailing economic conditions. In Eastern Canada the educational system seems to have been out of phase with the local environment for a long time. This has produced a colonial mentality and has oriented students into the wrong kinds of activities. Whether these claims can be justified in more careful and in more empirical studies than those contained in Ricker's collection is a moot point.

There are some difficulties with this volume which have been cited, but in large part it has served its purpose. To its credit, the work draws attention to the inadequacies of the schools and universities as they relate to the needs of people living in a non-urban environment. Essentially, this has been an urban century and education has tried to satisfy urban needs. The inadequacy of this approach, with related political implications, is fully revealed in Ricker's collection. Issues are raised, not settled; in doing this, the volume serves an important and useful function.

Essay Review II

MORAL EDUCATION: THE RISK OF OVERSIMPLIFICATION

Ivan DeFaveri
The University of Alberta

PREJUDICE. *Edited by L. Daniels, L. Douglas, C. Oliver, I. Wright, Association for Values Education and Research, University of British Columbia.* Toronto: The Ontario Institute for Studies in Education, Value Reasoning Series, 1978, 44 pp.; and accompanying TEACHER'S MANUAL, 32 pp.

THE ELDERLY. *Edited by L. Douglas, C. Oliver, I. Wright, Association for Values Education and Research, University of British Columbia.* Toronto: The Ontario Institute for Studies in Education, 1978, 40 pp.; and accompanying TEACHER'S MANUAL, 36 pp.

THE MORAL EDUCATION PROJECT (YEAR 5). *By C. Beck (Principal investigator), D. Boyd, E. Sullivan, J. Bradley, N. McCoy, S. Pagliuso.* Toronto: The Ontario Institute for Studies in Education, 1978, 180 pp.

Plato argued that the task of morally educating people was as difficult as, and inseparable from, the task of creating a just society. Many contemporary writers are now of the opinion that Plato made the problem more difficult than it really is, and we can now do with some ease what was once thought to be very difficult. We can now, it is alleged, by focusing on classroom activities—activities over which teachers have control, and without engaging in social and political reconstruction—do much more toward producing morally educated people than was hitherto thought possible, and this end result can be brought about without imposition, without manipulation and without indoctrination.

The task is complicated; the tendency to oversimplify must be resisted. Perhaps the most interesting oversimplification built into many of the contemporary programs of moral education is the claim that an adequate program of moral education need not concern itself with substantive principles because all that is really needed is the application of well-selected procedural principles. In short, it is sufficient to teach a methodology. This stance is a variation of the old slogan: do not teach students *what* to think; teach them *how* to think. Writers who follow this line of thought often invoke an analogy between the teaching of morality and the teaching of science: as it is possible to teach the scientific method without teaching specific scientific conclusions, so it is possible to teach a method appropriate to moral discourse while at the same time allowing the student to form his own conclusions. Moral education consists of teaching students how to be rational about morality, and surely no one could object to teaching students how to be rational. What could be more academically respectable than this? The task, then, is to teach a method and not to teach content, as the teaching of content would run too great a risk, there being no satisfactory way of deciding which content should be taught. The threat of indoc-

---

Dr. Ivan DeFaveri is Associate Professor in the Department of Educational Foundations at The University of Alberta. He received his Ph.D. at the University of California, Berkeley. His main interests are moral, aesthetic and religious values in education.

trinating or violating individual rights would be too great if content were allowed. The solution is not teach content at all. So it is argued.

I believe that this proceduralist account of moral education is seriously flawed, that a program of moral education can remain content-free only if it is simultaneously made trivial, and that any program of moral education that has the least importance must make reference to content. To insist on procedural principles alone is like saying to a student who wants to learn to play the piano that he should remember to play the right notes in the right order. In a sense, the advice is impeccable, but in isolation from some substantive comment about the correct order the notes are to take, it is of no significant value. Procedural principles that make no reference to content remain much too anaemic to form anything that could, without distortion of the language, be called a program of moral education. It is impossible to "do morality" in any important sense—to use an expression introduced by one of the writers in this school (Wilson, 1967, p. 26)—without doing a particular kind of morality, i.e., introducing substantive moral conclusions.

The programs that purport to restrict themselves to proceduralist principles rarely, if ever, exemplify the neutrality they advocate. Inevitably, they either smuggle in substantive values or, in various subtle and indirect ways, reinforce some of the values that are part of the social fabric in which they function.

Most of the English speaking writers in the area of moral education have been either English or American. We now have two Canadian organizations that have formulated programs of moral education designed for public schools. The Association of Values Education and Research (A.V.E.R.) at the University of British Columbia has recently published two small volumes, entitled *The Elderly* and *Prejudice*, and accompanying teacher's manuals. In Ontario, the Moral Education Project at the Ontario Institute for Studies in Education has recently published its final report describing the fifth year of its operation. In it, the reader is referred to publications by Beck for a more complete account of the program they are advocating.

The expressed aims of A.V.E.R. and the Moral Education Project are importantly similar. In the case of the former, we are told that their goal is to describe "how best to educate people in the domain of morality" (preface)[1]; in the case of the latter, "to help students reflect on their values . . . and . . . arrive at a sound set of values" (p. iii). But they have developed in individual ways and there are major differences between them. Both are at pains to preserve individual rights, to foster the autonomy of students and to avoid indoctrination. Both insist it is inappropriate to initiate students into any specific moral code. Each has made important contributions from which we can be guided in our attempts to influence the moral development of students, but both have serious limitation. Neither has entirely escaped variations of the proceduralist fallacy.

The Association of Values Education and Research

A.V.E.R. tells us that any example of sound reasoning about values takes the same form. Regardless of how complex, they assure us, "its basic logical structure remains the same" (p. 8). It takes the form of a syllogism. The major premise gives the "value standard" or moral principle that is

pertinent. The minor premise gives the relevant fact or facts. The conclusion, which must follow logically from the premises before it can be accepted, gives the value judgment applicable to the case at hand. They give an example:

Value premise (principle)	It is undesirable to poison streams
Factual premise	Strip mining in the Kootenays will poison streams
Value Conclusion	Strip mining in the Kootenays is undesirable

(p. 8)

Their aim, "to help people to learn to be rational about matters of morality" (p. 9), must be seen in conjunction with their belief that one is rational about moral matters if one uses the practical syllogism well. Of course, rationality demands choosing only relevant facts in the minor premise and only an acceptable moral principle in the major premise. But what counts as a relevant fact? And what counts as an acceptable value premise? Let us consider each in turn.

Unfortunately they have taken only minimal steps toward answering the first question. They let their argument rest on two examples, but it is impossible to extract from them any principle which may serve as an answer to the question.

The first example deals with a technical matter in elementary mechanics: "If I want to keep a large, heavy door from swinging shut but do not have a door stop," and I have a choice between a light pillow and a heavy stone, I would be irrational to choose the pillow (p. 4). Fair enough. If I need something heavy, and I have a choice between something heavy and something light, I should choose the heavy object. But whatever principle can be extracted from this example will not guide us in any important way if we are dealing with a genuine moral problem.

The second example describes a personnel manager who must hire someone for a position, the sole qualifications for which are competence and industriousness (p. 4). We are told that under these conditions, it would be irrational to consider the racial origin of the applicants. Of course. The previous decision that the "sole" qualifications are competence and industriousness make this conclusion inevitable—and uninteresting. But those people who find questions about race troublesome never have problems quite like this. In the example, the question whether race is relevant is not itself problematic, but it is precisely the question of whether or not race is relevant that defines the actual problems in some actual situations. Today, for example, the advocates and the opponents of affirmative action programs differ over precisely this point. Similar arguments abound in regard to the relevance of age in, for example, the amount of insurance premiums that a person ought to pay. Even in the example given, it is easy to imagine a different personnel manager, one who had yet to decide on the qualifications for the job, wondering if ability to get along with existing employees is an appropriate qualification, knowing that these employees were racial bigots. Here, he might on the one hand want to take advantage of the opportunity of civilizing them and make a point of

hiring someone from a different race, or on the other hand he might want to avoid the risk of decreasing their morale and lowering their productivity.

These examples tell us that what facts are relevant in any genuinely controversial matter is often itself a matter of dispute. Sometimes it is precisely this which distinguishes rival moral opinions. Insofar as the writers have raised the question of relevancy but have appealed only to unhelpful examples in their attempt to answer it, they have really avoided the question. I'm afraid some students may get the impression that it is easy in principle to decide which facts are relevant and thereby be left with the conclusion that making moral judgments is simpler than it really is.

Difficulties in deciding on the appropriate factual premise are of course paralleled by difficulties in choosing an appropriate value premise. In regard to this, the discussion is more detailed. They provide us with four tests designed to distinguish a major premise which is rationally acceptable from one which isn't. Any value premise which passes these tests must be seen as being as good as any other which passes the same tests. The tests, in summarized form, are:

1. *Role exchange test:* Ask if you would be willing to exchange places with the person(s) most disadvantaged in a situation. Imagine yourself in the circumstances of the most adversely affected person or group, then decide to accept or reject the value judgment and its action in the light of the consequences to the most adversely affected (p. 10).
2. *Universal consequences test:* Ask if you would be willing to accept a state of affairs where everyone acted according to the same principle (p. 10).
3. *New cases test:* Ask if you can accept all the judgments that issue from the principle in question. If there are legitimate exceptions, they should be built into a new principle (p. 11).
4. *Subsumption test:* Ask if your value principle follows logically (can be deduced) from a higher order value principle which is acceptable to you. If you cannot accept the higher-order principle, find new reasons for holding your original value principle, or reject it and reconsider your value judgment (pp. 11-13).

It is worthwhile noting just what these four tests accomplish, and what they do not. They are addressed to the person who has the motivation to be moral, but who wishes to be sure that the major premise of his practical syllogism is rationally justified. The tests, we are told, are for people who want to be "rational" about matters of morality, and who want to understand the "structure" and "logic" of morality (p. 8).

These tests have nothing to say to the person who sincerely wonders why he should enter the domain of rational morality in the first place, or who has no inclination to do so. They do not address themselves to the person whose motivation is such that the question of justification does not arise for him. Nothing here, in other words, is pertinent to the case of the egoist, who may understand the logic and the structure of morality as well as the next person, but who happens not to care about the well-being of anyone other than himself, and is quite willing to let others be egoists too. Similarly, the egoist may "imagine" the consequences for "the most adversely affected" people just as effectively as the next person, but if he doesn't care about the well-being of others, he may accept a principle that others are certain ought to be rejected. In light of the self-assigned task to see "how best to educate people in the domain of morality," it is difficult to see why they are

concentrating solely on the logic of justification and have so little to say about the psychology of people who may or may not want to use this logic. In the education of human beings, the education of the emotions cannot be so easily dismissed.

Let us leave the egoist and assume we are dealing with people who have the appropriate motives; that is, people who want to be moral, but don't quite know which "value premise" is morally acceptable. How much guidance would these four tests give him? Buddha's life changed, according to one legend, as a result of his imagining himself in the place of the aged and infirm. Many people, as a result of applying these tests consistently, would be persuaded to accept one value premise rather than another, and the world would probably be a more civil place if more people did apply them. People often judge their own conduct by a different set of rules from those by which they judge the conduct of others, and these tests are designed to thwart this tendency. But to say that many people, as a result of applying these tests and the recommendations implicit in them, will be *psychologically persuaded* to follow one path rather than another, is different from saying that a person applying them will be *logically compelled* to come to a particular conclusion. Given certain psychological dispositions, these tests would tend to rule out certain major premises and encourage others, but given a different set of dispositions, they could very well be used to support premises ordinarily found reprehensible. One can love others as much as he loves himself, but not love himself very much. The extent to which these tests assist in a program of moral education depends to a large extent on the kind of person who applies them, and how he describes the situation he is confronting.

Consider the role exchange test. Under certain conditions it could be seen as a very strong test, under other conditions very weak. If you took it very seriously, and in isolation from all considerations of desert or utility, it could be used to support a radical egalitarianism in which no one's total well-being was any worse than anyone else's. It could be used to argue that since no one, all else being equal, *is* willing to exchange places with anyone less well off than he is, then we ought to have a society where everyone's total good is identical. Only such a society would be a moral one. The differences tolerated by non-egoists are precisely those differences which they allow as a result of an appeal to their principle of justice, or their principle of utility, or because they believe such differences appropriately reflect innate differences. Without an appeal to any such putative justification for such differences, a radical egalitarianism such as this would be more rational than allowing the differences to continue. Viewed in this light, advocacy of this test can be seen as tantamount to advocacy of a particular substantive moral theory, presumably something the authors want to avoid.

But those who favour such a test rarely reach the conclusion that this kind of radical egalitarianism is appropriate. This is because it is rarely, if ever, used in isolation; it is almost always used in conjunction with principles of justice or utility, which it complements but does not supplant. If it is applied while simultaneously respecting certain moral principles having to do with desert or usefulness, it may alter nothing. There is a sense in which a person who restricts himself to moral considerations, and ignores

all other considerations such as those of prudence, aesthetics and religion, must be willing to exchange places with anyone who is treated morally. It is to this I now turn.

To say of someone that he is "adversely affected" or "disadvantaged" is to make a comment about the differences that exist between him and others. But these differences may be seen as morally acceptable or unacceptable. A person may lack certain material goods or opportunities, but it may be felt that his having them would not serve the common good or his having them is in some way deserved. In a hierarchical society such as Plato envisioned, for example, justice demanded that some have more than others. Usually, however, the expressions carry the suggestion that the disadvantaged person is unjustly disadvantaged and that the situation needs to be improved. Imagine a person who believes existing differences satisfy his sense of justice or his principle of utility, or who believes that in some way they are morally appropriate. He may believe, for example, that a person's well-being ought to be in direct proportion to his contribution to society. This person may be willing to consider the most adversely affected person, and to consider the possibility of exchanging places with him, but not without simultaneously asking whether each deserves his respective position or not. Of the person who is most adversely affected, he might say: if I contributed as little as he did, I would deserve as little as he is getting, but to receive less than I am now receiving while simultaneously continuing to make my present contribution would indeed be unjust. Nor would I exchange places with someone who is better off than I am if he deserves his better position; that also would be unjust.

It can be seen, therefore, that if a person has been treated justly, he cannot be said to have been disadvantaged in any moral sense. If his treatment has been morally satisfactory he can continue to be "disadvantaged" only in a nonmoral sense, with which this program is not concerned. Of such a person our hypothetical speaker may say that although he may not want to exchange places with him on, say, aesthetic grounds, he would be willing to change places with him on moral grounds, because both are being treated in a morally correct manner, and no one has the right to expect more than just treatment. Thus, such a person, that is, one who believes existing differences are justified would, if he restricted himself to normal considerations, be willing to change places with anyone. But it should be noted that his willingness to do this may be combined with a view of justice we find offensive. Thus the test may not accomplish what the advocates of moral education hope.

A second reason why this test may not do as much as its advocates hope becomes obvious once we realize that two people, both of whom may be willing to apply the test, may be unable to agree as to who actually is the most adversely affected person in a particular situation. To describe someone this way is not to offer a neutral description, but is almost always to offer a description from within a particular moral code. A person who holds a moral code different from ours may isolate a different person from the one we do, even though we both use the same words to describe our respective candidates. Thus, someone with a bizarre moral theory may acknowledge the legitimacy of the Role Exchange test, but his doing so

would not in any way improve the adequacy of his moral theory.

The status of the Universal Consequences test is in many ways similar. Many people will find in it a psychological persuasive reason to do one thing rather than another, but others won't. The test itself offers no logically compelling reason why students should accept any particular substantive conclusion. Some students may well appeal to this test as a way of legitimizing a conclusion others find repugnant, and do so on the grounds that since they are willing to have everyone follow the same principle, the principle ought to be tolerated. The example given as an illustration of the use of this test will make this clearer. Ted has cut across the grass:

> Ned: You shouldn't do that.
> Ted: Why not?
> Ned: Because you will wreck the grass.
> Ted: Phooey. How could I do any serious damage just by walking across it once?
> Ned: That may be true, but what if everybody walked across it?
> Ted: W-e-l-l, I suppose that would wreck it. But so what?
> Ned: It is easy to see what. If everybody walked across it, it would wreck the grass. And everybody should be treated fairly. That means we should all be treated alike unless we can show that there is something special about us (or our circumstances) that merits our being given special privileges. So unless you can show that you are special in a relevant way, you should stay off the grass. (p. 10)

Even in this simple example, one can see that this test carries weight only if there is previous agreement in regard to several substantive matters. There is no argument against the person who sincerely believes that the advantages of having a short cut outweigh the disadvantages of ruining the grass.

The implication left by the advocacy of this test is that any value principle that one is willing to universalize is morally acceptable. Now often this test will be sufficient to convince some people to modify their behaviour: they may realize that their actions should be judged by the same standards as they judge the actions of others, and thereby modify the behaviour that was about to make of them an unjustified exception. But in many cases of genuine moral dispute, it will have no force. Both sides of the abortion dispute are willing to universalize the principle that they see operative, but happen not to agree on what that principle is. Both are willing to universalize the principle that one ought not to murder, but the disagreement is about the metaphysical issue as to what counts as a human being. In regard to the same act, one would describe it as a case of murder and the other would not. The universal consequences test will not resolve the dispute, as both sides are willing to appeal to it. A pacifist and a militarist are both willing to live in a world in which everyone did what they were doing, but this in no way diminishes the differences between them. Egoists, mercenaries, amoralists, and fanatics are usually quite willing to have everyone follow the principles by which they live; often they are quite unhappy precisely because others don't.

The principle is effective if one who does not believe he has the right to think of himself as an exception finds himself doing just that. It will remind him that he ought to subject himself to the same rule he applies to others. Further, it may bring about a certain commonality of response among people who describe a moral problem the same way and who share a prior

commitment to certain substantive values. But it has no force against the person who is quite happy to see himself as unique and who applies unique standards to himself. Further, this principle, even if applied by everyone, is unlikely to increase agreement among people who sincerely describe moral problems differently.

In actuality these two tests, even when combined with the remaining two, do nothing but demand of the moral agent a pervasive consistency: he must judge identical acts the same way, he must prescribe to himself the same principle he prescribes to others, and the practical syllogisms he invokes must be internally consistent. But I think a person could successfully subject his major premises to each such test and still be the kind of person we would regard as morally miseducated. This "logic" in the hands of ordinary, decent people would allow them to be moral more consistently; but in the hands of the egoists, or the selfish, or those with odd emotional lives, or those with distorted theories of justice, it would be quite impotent. But it is precisely with these people that a program of moral education should be most concerned.

Throughout their discussion about "the logic of value reasoning," a position is taken in regard to the fact-value controversy, which is perhaps the issue more discussed in 20th century moral philosophy than any other. The position taken by A.V.E.R. is reflected in the centrality given to the practical syllogism and the process of deductive logic embedded in it. The message that the student will receive is that there is always a clear, unambiguous and easily defined difference between facts and values; that is, there is always a radical difference in kind between facts and values. An exercise given to the students will illustrate this. From a series of statements, students are asked to select those which are factual claims and those which are value judgments. For example, the statement, "If an ethnic group is unable to retain its language, then it loses its culture" is, we are told in the *Teacher's Manual,* a fact. But it would be just as correct to say, given the information that the student has, that this statement is nothing more than a definitional commitment, or that this statement emerges as a "fact" only if certain values are agreed upon, as the notion of "loss" and indeed the notion of "culture" are value-laden. Whether or not a sharp distinction can be made between normative and descriptive disciplines is deeply problematic, but the authors have settled for a particularly heavy-footed distinction, one deeply rooted in positivist thought throughout the centuries. This view has appeared, faded, and reappeared throughout history. I do not want to argue for or against the position taken by A.V.E.R. about this controversial matter, or whether Schopenhauer is right when he calls the position taken here "the greatest and most pernicious of errors, the fundamental error, the intrinsically *perverse* view" (1970, p. 133, italics in original). But I want to point out that I believe it is suspect, from the point of view of education, to pass off without discussion of its logical status, something which is really problematic. We are told that "this program proceeds on the assumption that there is a rational approach to the solution of moral problems" (p. 5). Well and good; but the further, unstated assumption is that there is only one rational solution to the fact-value problem. Some students will not recognize this as a further assumption; indeed, they will believe it is a "fact" that their world is divided into two

discrete and mutually exclusive categories, and that value-free nature contrasts sharply with man's imposed evaluations.

The Moral Education Project

The publication of the Moral Education Project at O.I.S.E. here considered describes the fifth year of their operation. It begins with a summary of the Project's "reflective approach" to values education, the justification of which is given in greater detail in other publications (specifically, Beck, 1971, 1972, 1974, 1976). The program is also referred to as the "fundamental human values" of "ultimate life goals" approach because of the pivotal place ultimate values are seen to occupy. The function of a moral education program is to show people the best way to attain their ultimate values. Of ultimate values themselves, three claims are made. First, human beings tend to pursue the same ultimate values. They tell us that "as we become aware of the connections between our life goals, we will notice that there is more in common amongst the life goals of people around the world than we had previously thought" (Beck, 1972, p. 10). Great commonality among human beings in regard to ultimate goals is envisioned:

> The great majority of human beings pursue ultimate life goals, or fundamental human values such as survival, happiness (enjoyment, pleasure, etc.), health, fellowship, (friendship, love, etc.), helping others (to some extent), wisdom, fulfillment (of our capacities), freedom, self-respect, respects from others, a sense of meaning in life, and so on. (Beck, 1976, p. 1)

Second, each ultimate life goal must be given equal status. "The old notion of a hierarchy of values . . . is quite unworkable . . . . It is not possible to develop a ranking of goals such that one can simply read off which to pursue when two goals conflict" (Beck, 1976, p. 269). If they do conflict, he tells us, unhelpfully, that, "appeal must be made to the maximization of life goals, in the long run and on the whole and all things considered" (Ibid, p. 269). Third, ultimate life goals cannot be subjected to rational scrutiny. People "just *do* pursue them" (Beck, 1974, p. 271, italics his). Rationality enters only as a judgment on the means chosen to attain these ends, but not as a judgment on the ends themselves. He is quite adamant about this. It is "meaningless", he insists, to ask whether people in general ought to pursue the goals they actually do pursue (Beck, 1974, p. 24). Elsewhere we are told that he is talking about "mature" human beings only, and it is necessary to reject the choices made out of ignorance, prejudice and superstition (Ibid, p. 25). There are a number of problems with this view of the valuing process. To say that the great majority of human beings tend to want some or all of the ultimate life goals listed here is misleading. At this level of abstraction, indeed almost all people will tend to use the same words to describe what they are after, but the agreement may be purely verbal. It is all very well to say that both Arabs and Israelis want justice—indeed they do—but the sense in which they want the same thing can be seen as trivial, especially at the onset of each war. Both the Marquis de Sade and his victims wanted happiness. Both capitalists and socialists strive for the fulfillment of human beings. Such terms serve a purpose, but in isolation they do not tell us anything important. Only if the context is made clear, or if the statement of the ultimate goal is followed by a series of progressively more specific

statements about the goal, will it be sufficiently informative in any reasoned account of how we should govern our moral lives.

The claim that people just do accept certain goals as ultimate, and that the task of moral education is merely to give them the ability to attain these goals efficiently is a philosophically respectable position, but there are other conceptions about the nature of ultimate goals that need to be taken seriously and about which a teacher genuinely interested in education would insist on his students knowing.

The orientation reflected here is embedded in the empiricist tradition and is sometimes encapsulated in the slogan, "good reasons apply to means, not ends." Any end is morally acceptable as long as attaining it does not interfere with the pursuit of a more valued end. The ancient tension between what men "just do" in regard to the choosing of ultimate ends, and what they should do—either because of their own nature, or because of the structure of the world—dissolves here. This view, embedded in the ultimate life goals approach, contrasts with various intellectual traditions which claim that even ends can be judged rationally, and that man, not entirely immanent in time, can in some important sense transcend himself. The former view sees morality as an applied social science and the making of moral decisions essentially as a technical matter. The latter view believes that morality, properly conceived, reflects an independent, critical, and autonomous reason and rejects the view that reason is purely instrumental. This continuing debate in intellectual history and philosophy is far from over. As Kolakowski reminds us, "philosophy has never given up its attempt to constitute an autonomous 'reason' independent of technological applications" (1969, p. 210). But this highly problematic matter is here treated as if there were no dispute; insofar as this is done, the purposes of moral education are poorly served. If we are interested in educating students in the domain of morality, rather than just encouraging them to adopt a particular moral theory, we should make sure that they are familiar with the most important controversies that are part of its evolving history.

Even with the conception of ends that is entertained here, it is still odd to say that people "just do" choose ultimate ends, especially in an educational context. For people choose one end rather than another in part because of certain experiences they have had and certain influences that are exerted upon them. If teachers don't encourage students to choose one set of goals rather than another, it will be done by default by advertisers and con men of various persuasions. Further, the list of fundamental life goals given here reflects a curiously optimistic view about the nature of man. It is as if all his fundamental impulses were good, and there were no darkness in his heart. But many people, I suspect, if they were not self-deceived and if social conditions allowed, would truthfully tell us that a list of their ultimate goals would include such things as self-aggrandizement, mastery over others, and a sadistic impulse to destroy. It is on the periphery of this point that Beck makes a startling admission, but one the significance of which he seems to be unaware. In regard to the "reflective approach," he says, "In order to reflect successfully on values, a student must already have arrived at a relatively sound value outlook" (1970, p. 3). It seems that the approach only works with students who need it least, and nothing has been said about

what is to be done with those students who need moral education most. In our school system, it is not the student who has already arrived at "a relatively sound value outlook" who is bothering people; it is the student who is the egoist, or who is insensitive to the needs and feelings of others, or who is not influenced by moral arguments.

As was suggested earlier, no attempt has been made to evolve a principle which can settle conflicts, either conflicts which arise because a person holds mutually exclusive ends, or one which arises between two people. We are told only that these people "should be shown that the major concern of ethics is to try to arrive at mutually beneficial arrangements or 'happy compromises' " (Beck, 1971, p. 33). This is advice which every high school girl with an unwanted pregnancy should read very carefully.

Conclusion

The problems that arise when an attempt is made to initiate a program of moral education are very difficult in a country which celebrates cultural diversity, respects individual rights and personal autonomy, believes education is something beyond socialization, and insists on compulsory education. Perhaps it is impossible to have a program of moral education without modifying some of the good things simultaneously wanted. The two Canadian programs represent bold attempts to satisfy the considerable demand for practical programs in this sensitive area. Insofar as they will encourage students to realize that it is often inappropriate to hold values at a pre-critical and pre-reflective level, and that it is often necessary to display public reasons for one's moral conclusions, they will have done a great service. But this promise of future success must be tempered with an understanding of the limitations the programs themselves impose.

In regard to A.V.E.R., one is struck with the incongruity between their reluctance to enter the area of content and the forthrightness with which they commit themselves to a particular, controversial view of moral reasoning. As I have argued, I believe the first step here is to realize that a program of moral education which restricts itself to procedural principles alone will accomplish less than what most people expect from it. It is to the credit of the Moral Education Project at O.I.S.E. that they have seen the need to discuss substantive values. But in their assumption that, after all, everyone's goals are similar and that being moral is nothing more than the efficient attainment of these goals, they have treated a long-standing dispute in the history of thought as if it had been finally settled. In this regard, perhaps neither program takes the notion of education sufficiently seriously. Students guided by such programs could quite easily be misled about the status of certain claims made by the authors. They might not realize that the positions taken in regard to, for example, the nature of morality, the scope of deductive logic within moral discourse, the adequacy of the means-end model in this area, and the nature of rationality, are themselves quite controversial, and by no means the only positions entertained by serious writers in moral philosophy. I find this an unhappy situation, because I believe that although a teacher in a public school need not be impartial in regard to controversial matters, he ought to represent competing positions in a fair manner. By not telling the students about

alternative accounts, the writers are being less than attentive to the desirability of promoting autonomous students, an ideal to which they implicitly subscribe.

Thus, these programs restrict themselves to a limited series of questions. A more complete investigation of what it is to be morally educated would include a study of the forces which help shape what one of these programs refers to as our ultimate goals and the other program as our willingness to live by certain principles. This would incorporate, at least, a study of insights from psychology, history, and sociology. All the programs mentioned in this essay reflect an individualist stance. But if we believe that, in part, a person's self is an historical and social product, we will not be satisfied with a program of moral education until it considers the possibility of influencing the factors which help form the individuals involved. Thereby, we will realize that the problem of creating moral people must be seen, at least in part, as the problem of social justice. Similarly, a program of moral education ought not to overestimate the value of instrumental reason. Perhaps rationality, in this sense, is not all we need. G. J. Warnock has written that if people choose not to want to be moral, "the essence of the remedy is not in reason, it is in non-indifference" (1973, p. 166). One can "reason" and "reflect" but be insensitive to the suffering of others. What is it that will bring the capacity to care about others into being, and make a person realize that the pain he himself suffers is no more important than the pain that others suffer? An instrumental view of reason will not help here, because according to this conception, the notion of reasonableness arises in the area of morality only if such a commitment is presupposed; to decide to care about others would be a pre-rational or nonrational commitment. But if one doesn't care about others, neither talk about "logic" nor a systematic listing of actual desires will do much good. Some, as was noted earlier, will really want to pursue destructive ends regardless of what they profess. We may discover that in the final analysis, being moral involves more than having mastered a technique.

Notes

1. All questions from the A.V.E.R. material are in *Prejudice: Teacher's Manual.* The same quotations appear in *The Elderly: Teacher's Manual,* but the pagination is different.

References

Association for Values Education and Research. *Prejudice,* ed. by L. Daniels, L. Douglas, C. Oliver, I. Wright. Value Reasoning Series, The Ontario Institute for Studies in Education, 1978, and accompanying *Teacher's Manual.*

Association for Values Education and Research. *The elderly,* ed. by L. Douglas, C. Oliver, & I. Wright. Value Reasoning Series, The Ontario Institute for Studies in Education, 1978, and accompanying *Teacher's Manual.*

Beck, C. *Moral education in the schools.* Toronto: The Ontario Institute for Studies in Education, 1971.

Beck, C. *Ethics.* Toronto: McGraw-Hill, 1972.

Beck, C. *Educational philosophy and theory.* Boston: Little, Brown and Company, 1974.

Beck, C. *The reflective approach in values education.* Toronto: The Ontario Institute for Studies in Education, 1976.

Kolakowski, L. *The alienation of reason: A history of positivist thought.* New York: Doubleday, 1969.

The Moral Education Project (Year 5), C. Beck (Principal Investigator), D. Boyd, E. Sullivan, J. Bradley, N. McCoy, S. Pagliuso. Ontario Ministry of Education, The Ontario Institute for Studies in Education, 1978.

Schopenhauer, A. *Essays and aphorisms.* Harmondsworth, Middlesex: Penguin, 1970.

Warnock, G. J. *The object of morality.* London: Methuen, 1973.

Wilson, J. *Introduction to moral education.* Harmondsworth, Middlesex: Penguin, 1967.

BOOK REVIEWS

CLASSROOM LANGUAGE: WHAT SORT? *By Jill Richards*. Winchester, Mass.: Allen & Unwin Inc., 1978, 154 pp.

Classroom Language: What Sort? explores issues related to language learning and learning through language in schools. Professor Richards introduces this exploration with reference to what she identifies as a current "dilemma for schools." This dilemma is seen as the result of pressures to educate for the demands of modern technological society coupled with the concern for providing opportunities for personal growth and development. The implication of these competing demands for classroom language are discussed with reference to present practice and to the thinking and research of prominent educators and linguists. Professor Richards' own research in British schools provides input to the six-chapter text which challenges some popular beliefs related to school practice and urges some rethinking about the nature and scope of school language programs.

If there is a fundamental argument guiding the discussion in the text, it is a plea to resist simple answers or solutions in treating such complex phenomena as language learning and language itself. What are described as "language-limiting" practices and "class" stereotyping of students are cited as prime examples of simple and inadequate responses to complex problems and issues. It is argued, for example, that many so-called "child-centered" approaches to language learning in the early school years are, in fact, inadequate with respect to language function emphasis and too restrictive in terms of the experiences provided for students. Various practices associated with developing "self-expression" and "creativity" are cited as examples of focusing on "poetic function" to the exclusion of other language functions, and emphasizing "first-hand" experiences to the exclusion of other viable alternatives in providing opportunities to learn and use language.

With respect to the language learner or "user," Richards argues against the use of social class stereotypes in explaining and predicting language behaviour. Stating that much of the sociolinguistic research has produced conflicting conclusions, she reviews the work of Bernstein ("codes") and Labov ("dialect") in particular. Again her argument rests on the criterion of adequacy. She claims that the theories of language learning and development restricted to social class influences are inadequate explanations of complex phenomena. She argues further that we will come no closer to understanding language behaviour until we look at a number of factors influencing language within individual, rather than group, life-style. She

notes, also, that many children whose language does not represent "code" or "dialect" have difficulty meeting some of the language demands placed on them. Richards states that this is particularly evident in what is described as the transition from primary to secondary school.[1]

The kinds of problems associated with the language and thinking demands of programs in the later school years are identified in discussions of "language demands of subject learning" and "specialized language and concept development." Although some of the difficulties experienced by older students are traced to the "language-limiting" nature of primary level programs, Richards does not excuse subject area teachers from any responsibility in this regard. Much of the discussion related to language demands in subject areas centers on Richards' research in classrooms. The research was designed to test hypotheses related to children's "ordinary language" and the language encountered in school, different language demands of different subject areas, and the linguistic features of the language in different subject areas. The reports of Richards' investigation and the work of other researchers identify the conceptual tools and systems of analysis used in this kind of research on classroom language. The findings of this research, and the discussion of specialized language and concept development related to subject learning, raise questions which have implications for teachers' language and instructional strategies and techniques. The main thrust of these discussions appears to be an attempt to develop awareness of the role of language in learning across the curriculum. Implicit also is the notion that the school language program needs to encompass all facets of school life and learning.

The "conclusions and recommendations" set forth in the final chapter suggest how teachers might make use of the text's discussions in examining language program and practice. Pertinent recommendations from the Bullock Report (1974) are reviewed also in relation to issues raised throughout the text.

Professor Richards has identified and analyzed a number of significant issues in a relatively slim volume of 144 pages. Although there is an attempt to link discussion throughout the text, chapters 2, 3, 4, and 5 are, in a sense, self-contained with respect to focus and argument. Thus the text is best described, perhaps, as a series of discussions on different issues in language education. However, this feature of the text suggests its value as a reference or resource work. With one exception, the arguments in the text are clearly developed and not overstated. The exception is the attempt to challenge the conclusions of some sociolinguistic studies. In this discussion and analysis, Professor Richards comes close to arguing herself into a "semantic" corner. Nevertheless, the text stands as a scholarly piece of work which should be of interest to teachers and to all concerned with language programs in schools.

Notes

1. The designations of "primary" and "secondary" school refer to the organizational structure in England. Children of 12 years are identified as being in the first years of secondary school.

Ellen M. Regan
Department of Curriculum
The Ontario Institute for Studies in Education

AN EXPERIMENT IN TEACHER EDUCATION. *By Alec Ross, David McNamara and Jean Whittaker.* Guildford, Surrey: Society for Research into Higher Education, Ltd., 1977, 133 pp. (ISBN 0-900868-554)

This monograph is based upon the final report of an experiment in teacher training which was undertaken in the northwest of England in the period 1971-1974. It involved some fifty-two students in three small colleges associated with the University of Lancaster. The main thrust of the experiment was to substitute courses, in which the subject matter of several disciplines taught in schools was "applied" to the education of children, for courses of considerable depth in one or two academic disciplines. This concept is in line with the thinking of the Bicentennial Commission on Education for the Profession of Teaching which completed its work in the United States in 1976. This Commission recommended that principles be the "constants" of academic study rather than have students concentrate on a maze of specific topics. The logic behind this substitution of courses, or, in the American context the emphasis on principles rather than on the details of a discipline, is straightforward. It is that teacher education should be designed to help teachers teach youngsters the basic elements of the disciplines rather than to aspire to make prospective teachers into mathematicians, scientists or whatever.

The actual "Applied Education Course" attempted to cover the subject content appropriate to the learning of children of the 5-13 age group. Modules concerned with key areas of the schools' curricula were developed. Two modules, English and Mathematical Education, were compulsory with students given choices among several others. The overall Applied Education package involved considerably more than merely the teaching of subject-matter but included the development of attitudes and skills. These additional objectives may have had some confounding effects on the experiment.

The description of the experiment is exceptionally detailed and rather elaborately presented so that the reader tends to get an exaggerated notion of its significance or sophistication. The principle involved is an important one, but a relatively small sample, together with a rather hurriedly conceived and perhaps amateurishly implemented test, should not persuade too many to rush to follow suit. The findings, for what they are worth, do seem to indicate that such an approach tends to be better preparation for teaching in "primary" schools than is the traditional in-depth subject matter route. College staff appeared to welcome it and students were judged to be more enthusiastic. There was no indication that such a course failed to "stretch" intellectually the experimental group. In fact, those of the experimental group who continued their professional education to the B.Ed. level did particularly well.

The general impression conveyed by this account of the experiment is that, for students firmly committed to careers in primary teaching, study of the application of discipline content to teaching is satisfying and appropriate. For students lacking this firm commitment such an approach has less value. On balance it appears to this reviewer that the monograph is a useful addition to the literature concerned with teacher education but that its substance could have been conveyed in considerably fewer pages.

Lorne D. Stewart
Department of Secondary Education
The University of Alberta

THE ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

Volume XXV, 1979—Table of Contents

Adams, B. B., and Ollila, L. O. *The relationships of language concepts, auditory comprehension, visual perception, and spatial relations as predictors to reading achievement in first grade.* No. 4, December, 248-258.

Aikenhead, G. S. *Using qualitative data in formative evaluation.* No. 2, June, 117-129.

Assheton-Smith, M., and Toohey, K. *School-centered community conflict: The Holdeman Mennonite case in Alberta.* No. 2, June, 77-88.

Bacchus, M. K. *Education as a social control mechanism.* No. 3, September, 160-173.

Bragman, R., and Hardy, R. *Identical and reverse visual pattern recognition in deaf children.* No. 3, September, 174-181.

DeFaveri, I. *Moral education: The risk of oversimplification* (Essay Review). No. 4, December, 294-306.

Evans, R. N. *Some observations regarding value added by education* (Perspective). No. 4, December, 284-287.

Hamalian, A. *Pupil control ideology: Comparative Perspectives—United States and Canada.* No. 1, March, 37-47.

Horodezky, B. *Comparative difficulty of beginning reading vocabulary: Set II.* No. 4, December, 259-263.

Hughes, A. S. *Curricular knowledge organization and variations in instructional emphases.* No. 1, March, 4-19.

Kehoe, J. W., and Ungerleider, C. *The effects of role exchange questioning on empathetic perceptiveness.* No. 1, March, 48-52.

Kieren, D. K. *The child and educational research—For whose benefit?* (Guest editorial). No. 1, March, 1-3.

Kieren, T. E., and Southwell, B. *The development in children and adolescents of the construct of rational numbers as operators.* No. 4, December, 234-247.

Kirman, J. M., and Goldberg, J. *Student teacher telephone conferencing with satellite maps as a monitoring device.* No. 4, December, 275-283.

Liedtke, W., and Stott, G. *Games and game settings for the preschool child.* No. 3, September, 182-191.

MacIver, D. A. *Education in Atlantic Canada* (Essay Review). No. 4, December, 288-293.

Mason, J. M., and Kendall, J. R. *Facilitating reading comprehension through text structure manipulation.* No. 2, June, 68-76.

McArthur, J. *Teacher socialization: The first five years.* No. 4, December, 264-274.

Nwankwo, J. I., and Ohikhena, T. O. *Perception of the administrative role of the principal as a factor in student conflict in Nigeria.* No. 3, September, 147-159.

Panuto, B., and White, D. *Achievement and the prediction of achievement in English first and second language children.* No. 2, June, 61-67.

Pereira-Mendoza, L. *Heuristic strategies utilized by high school students.* No. 4, December, 213-220.

Schlosser, L., and Algozzine, B. *The disturbing child: He or she?* No. 1, March, 30-36.

Schnell, R. L. *History of childhood as history of education: A review of approaches and sources* (Perspective). No. 3, September, 192-203.

Shaw, E. *The effects of attentional focus on graphic discrimination.* No. 2, June, 103-116.

Travis, L. D. and White, W. B. *Experimental manipulation of the recall of narrative material by five-year-olds.* No. 3, September, 137-146.

Walker, L. *Newfoundland dialect interference in fourth grade spelling.* No. 4, December, 221-233.

Williams, D. M. *A study of moral education in Surrey, B.C. secondary schools.* No. 2, June, 89-102.

Young, J. H. *The curriculum decision-making preferences of school personnel.* No. 1, March, 20-29.

PUBLICATIONS

FACULTY OF EDUCATION, UNIVERSITY OF ALBERTA EDMONTON T6G 2G5

The Alberta Journal of Educational Research

AJER is a quarterly journal devoted to the dissemination, criticism, interpretation and encouragement of all forms of systematic enquiry into education and fields related to or associated with education. Published in March, June, September and December; subscription is $8 per year. Address communications to the Editor, *AJER,* 732 Education South.

Elements

Elements endeavours to help educators translate theory and findings from current research into classroom practice. Each issue focuses on a specific subject area, topic or theme. Published monthly September through April; subscription is $4 per year. Address communications to the Editor, *Elements,* Department of Elementary Education, 539 Education South.

Indian-Ed

Indian-Ed publishes articles, presents abstracts of research and reports news items related to the education of native peoples. Published four times per year; subscription is $3. Address communications to the Editor, *Indian-Ed,* Department of Educational Foundations, 5-109 Education North.

The Canadian Administrator

The **CA** serves administrators at all levels of educational organizations. Articles selected for publication are intended to familiarize practitioners with the findings and implications of research, and with analyses of current policies and practices. Published eight times per year October through May; subscription is $5 per year. Address communications to Editor, *CA,* Department of Educational Administration, 7-104 Education North.

CONTEMPORARY EDUCATION

The NOW Journal of Teaching and Learning

CE: Publishes articles by teachers, professors and school administrators
CE: Reviews current books in the broad field of education
CE: Helps educators with their professional tasks
CE: Offers lively editorials
CE: Gives you regular columns on provocative topics, people
a. A Turning Point*
b. Meandering— M. Dale Baughman**

*Readers describe special moments in education.
**Professor of Secondary Education, and Editor of CONTEMPORARY EDUCATION.

CONTEMPORARY EDUCATION, Statesman Towers
Room 1005, Indiana State University
Terre Haute, Indiana 47809

Please send the next 4 issues of CONTEMPORARY EDUCATION for $7.00 ($10.00 for foreign countries).
_____ Payment enclosed _____ Bill me

Name _____

Address _____

City _____

State _____ Zip _____

PREPARATION OF MANUSCRIPTS

1. All manuscripts must be typewritten, double spaced, and submitted in duplicate. An abstract of approximately 100 words in length, typed on a separate page, should be provided.
2. Tables must be numbered in Arabic numerals with the word 'Table' centered and in capital letters, e.g., TABLE 4. The heading of the table is to be centered below and typed in capitals. The format of tables should conform to the specifications in the APA Publications Manual.
3. Graphs and charts should be used only if essential. They must be carefully prepared on separate sheets in India ink, ready for reproduction. Graphs must be properly labelled using Arabic numerals, e.g., Figure 3.
4. Each table or figure should be presented on a separate page. The position of tables and graphs should be clearly indicated within the text by inserting at the relevant point the phrase (Insert Table 2 here).
5. References should appear in parentheses following the reference citing the author's name (unless the name appears in the text), the year of publication, and page number if appropriate. For direct quotations, the reference should be cited and the page number given in brackets before the final punctuation of the quotation. The references should be listed alphabetically by author's last names at the end of the manuscript under the heading, *References*.
6. Explanatory notes, numbered consecutively and identified in the text with a superscript, may be included under the heading of *Notes*. They should be double spaced and placed at the end of the manuscript immediately preceding the *References*. The citing of references and quotations in the *Notes* should conform to the procedures outlined in No. 5 above.
7. Spelling shall conform to the *Oxford English Dictionary,* except in the case of direct quotations which must conform to the original. Editorial alterations will be made if necessary.
8. In matters of style, the APA Publications Manual is considered definitive.

PRINTED BY
THE UNIVERSITY OF ALBERTA
PRINTING DEPARTMENT